

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 10:44:18 ; Search time 3858 Seconds
(without alignments)
10248.693 Million cell updates/sec

Title: US-10-720-018-1

Sequence: 1 atgctgactcattccacac.....atggtgttcactcttgg 816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmb1.*

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_ay:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 816 | 100.0 | 816 | AJ583159 | AJ583159 Ambrosioz |
| 2 | 332.8 | 40.8 | 1166 | CTARDH | U00675 Candida tro |
| 3 | 305.6 | 37.5 | 1614 | YGAARDH | L16227 Candida alb |
| 4 | 275.2 | 33.7 | 110000 | CR382133_10 | Continuation (11 o |
| 5 | 268.4 | 32.9 | 1157 | PSARDH | Z46866 P. scriptis |
| 6 | 238.2 | 29.2 | 696 | AR345028 | AR345028 Sequence |
| 7 | 238.2 | 29.2 | 696 | AR345028 | AR345028 Sequence |
| 8 | 218.6 | 26.8 | 846 | CR382121_09 | Continuation (10 o |
| 9 | 155.8 | 19.1 | 110000 | CR382133_03 | Continuation (4 of |
| 10 | 130.8 | 16.0 | 110000 | CR382133_11 | Continuation (12 o |
| 11 | 92.6 | 11.3 | 110000 | CR382133_06 | Continuation (7 of |
| 12 | 91.4 | 11.2 | 110000 | CR382133_06 | Continuation (7 of |
| 13 | 87.2 | 10.7 | 4293 | AF002134 | AF002134 Candida a |
| 14 | 86.8 | 10.6 | 126105 | AC105434 | AC105434 Magnapor |
| 15 | 86.6 | 10.6 | 927 | AR547359 | AR547359 Sequence |
| 16 | 81.8 | 10.0 | 1485 | AY387591 | AY387591 Emericell |
| 17 | 81.8 | 10.0 | 15779 | AY434691 | AY434691 Emericell |
| 18 | 78.6 | 9.6 | 110000 | CR382125_21 | Continuation (22 o |
| 19 | 78.6 | 9.6 | 110000 | CR382131_06 | Continuation (7 of |

| | | | | | | | |
|---|----|------|-----|--------|----|-------------|-----------------------|
| C | 20 | 72 | 8.8 | 110000 | 1 | AE016827_14 | Continuation (15 o |
| C | 21 | 70.8 | 8.7 | 110000 | 8 | CR382133_10 | Continuation (11 o |
| C | 22 | 70 | 8.6 | 1063 | 8 | UF081790 | U01790 Uromyces fa |
| C | 23 | 68.6 | 8.4 | 409 | 11 | KLAJ9366 | AL229366 Kluverom |
| C | 24 | 68.2 | 8.4 | 340900 | 1 | SME591791 | AL591791 S. pombe chr |
| C | 25 | 67.6 | 8.3 | 32814 | 8 | SPAC22A12 | Z99295 S. pombe chr |
| C | 26 | 64 | 7.8 | 6430 | 8 | ARJ320167 | AJ320167 Aspergill |
| C | 27 | 64 | 7.8 | 110000 | 8 | CR382138_00 | CR382138 Debaryomy |
| C | 28 | 63.6 | 7.8 | 10126 | 1 | AE001711 | AE001711 Thermotog |
| C | 29 | 62.4 | 7.6 | 309135 | 1 | AP005946 | AP005946 Bradyrhiz |
| C | 30 | 61.8 | 7.6 | 873 | 6 | AR547523 | AR547523 Sequence |
| C | 31 | 61.4 | 7.5 | 110000 | 6 | CR382132_02 | Continuation (3 of |
| C | 32 | 61 | 7.5 | 786 | 8 | STU245634 | AJ245634 Solanum t |
| C | 33 | 59.8 | 7.3 | 822 | 6 | AX772916 | AX772916 Sequence |
| C | 34 | 59.8 | 7.3 | 1129 | 8 | DURTROPR1 | L20473 Datura stra |
| C | 35 | 59.4 | 7.3 | 15350 | 8 | SPAC8B11 | AL021819 S. pombe c |
| C | 36 | 59 | 7.2 | 774 | 6 | E30144 | E30144 Arabiditol |
| C | 37 | 59 | 7.2 | 2255 | 6 | E30145 | E30145 Arabiditol |
| C | 38 | 57.8 | 7.1 | 300787 | 1 | AE016744 | AE016744 Staphyloc |
| C | 39 | 57.2 | 7.0 | 37969 | 8 | CR382137_20 | Continuation (21 o |
| C | 40 | 57.2 | 7.0 | 110000 | 8 | CR382137_19 | Continuation (20 o |
| C | 41 | 56.6 | 6.9 | 976 | 8 | STU400815 | AJ400815 Solanum t |
| C | 42 | 56.4 | 6.9 | 789 | 8 | BT010425 | BT010425 Arabidops |
| C | 43 | 56.4 | 6.9 | 825 | 6 | AX772918 | AX772918 Sequence |
| C | 44 | 56.4 | 6.9 | 909 | 8 | AKI75221 | AKI75221 Arabidops |
| C | 45 | 56.4 | 6.9 | 1099 | 8 | D88156 | D88156 Hyoscyamus |

ALIGNMENTS

| | | | | | |
|------------|--|--------|---|--------|-----------------|
| RESULT 1 | AJ583159 | 816 bp | mRNA | linear | PLN 13-APR-2004 |
| LOCUS | AJ583159 | | Ambrosiozyma monospora partial mRNA for NADH L-xylose reductase | | |
| DEFINITION | AJ583159 | | (Alx1 gene). | | |
| ACCESSION | AJ583159 | | GI:42820317 | | |
| VERSION | AJ583159.1 | | Alx1 gene; NADH L-xylose reductase. | | |
| KEYWORDS | Ambrosiozyma monospora | | | | |
| SOURCE | Ambrosiozyma monospora | | | | |
| ORGANISM | Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycopsidaceae; Ambrosiozyma. | | | | |
| REFERENCE | Verho, R., Putkonen, M., Jondesborough, J., Penttila, M. and Richard, P. | | | | |
| AUTHORS | J. Biol. Chem. 279 (15), 14746-14751 (2004) | | | | |
| TITLE | Carabolic Pathway of Yeast | | | | |
| REFERENCE | Richard, P. | | | | |
| AUTHORS | Submitted (18-SEP-2003) Richard P., VTT Biotechnology, VTT, PO box 1500, Tietoele 2, Espoo, 02044 VTT, FINLAND | | | | |
| TITLE | Location/Qualifiers | | | | |
| JOURNAL | 1. 816 | | | | |
| source | /organism="Ambrosiozyma monospora" | | | | |
| FEATURES | /mol_type="mRNA" | | | | |
| gene | /strain="NRRL Y-1484" | | | | |
| CDS | /db_xref="taxon:43982" | | | | |
| | 1. 816 | | | | |
| | /gene="alx1" | | | | |
| | 1. >816 | | | | |
| | /gene="alx1" | | | | |
| | /codon_start=1 | | | | |
| | /product="NADH L-xylose reductase" | | | | |
| | /protein_id="CAE47547.1" | | | | |
| | /db_xref="GI:42820318" | | | | |
| | /translation="MTDYLPTFPEDGHLITVGCAGGALALIKGLAYGSDIALDI | | | | |
| | DOEKTAQKAEHYKATVETELKKEVPMGSAVDISPDVHVFPQVAAADPKPLPH | | | | |
| | LVNTAGCCEPCEPDYPAKAEKMGVKNVNLGSLIVSAPAKPLIKSGIKASVYLIGS | | | | |
| | MSGATVNDPQNVNMSKAGVTHLAKTLACVAKTNRIVNSLNPGITGPLTKNVIN | | | | |
| | GNELYNRWISGIPIQORMSEPKYIGAVIVYLSESAASYTTGASLLVDGGFTSW" | | | | |

| | |
|-----------------------|---|
| ORIGIN | |
| Query Match | 100.0%; Score 816; DB 8; Length 816; |
| Best Local Similarity | 100.0%; Pred. No. 6.3e-196; |
| Matches | 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| QY | 1 ATGACTGATCACTTCCAACTTTAGATGATGATGAGCCCTTACCATGTCACAGGTC 60 |
| DB | 1 ATGACTGATCACTTCCAACTTTAGATGATGATGAGCCCTTACCATGTCACAGGTC 60 |
| QY | 61 TGTGGTGGTTAGCTGAAGCTTTATCAAGGGTTTGTGGCTTACGGTTCTGACATTT 120 |
| DB | 61 TGTGGTGGTTAGCTGAAGCTTTATCAAGGGTTTGTGGCTTACGGTTCTGACATTT 120 |
| QY | 121 TTGCTTGATTCGACCAAGAAAAGACTGCTCCCAACAGCGAATACCAAAATAGCCT 180 |
| DB | 121 TTGCTTGATTCGACCAAGAAAAGACTGCTCCCAACAGCGAATACCAAAATAGCCT 180 |
| QY | 181 ACTGAAAGATGGAAGTTGAAAGAGTTCCAAAGATGGGTTGATATGCTGTGATTTCT 240 |
| DB | 181 ACTGAAAGATGGAAGTTGAAAGAGTTCCAAAGATGGGTTGATATGCTGTGATTTCT 240 |
| QY | 241 GATTCTGATACCGTTCAACAAGGTGTTGCTCAAGTTGCTAAGATTTTGTAAGTTGCA 300 |
| DB | 241 GATTCTGATACCGTTCAACAAGGTGTTGCTCAAGTTGCTAAGATTTTGTAAGTTGCA 300 |
| QY | 301 TTGCACTTGTTAACAACAGCTGCTGTTACTGTGAAAACCTTCCATGTGAAGATTTACCCAGCC 360 |
| DB | 301 TTGCACTTGTTAACAACAGCTGCTGTTACTGTGAAAACCTTCCATGTGAAGATTTACCCAGCC 360 |
| QY | 361 AAGAACGCTGGAAGATGTTGAAGGTTAACTTGTGGGTTCTTTGATGTTCTCAAGCC 420 |
| DB | 361 AAGAACGCTGGAAGATGTTGAAGGTTAACTTGTGGGTTCTTTGATGTTCTCAAGCC 420 |
| QY | 421 TTGCTAAGCCATTTGATCAAGAGGTTATCAAGGGTCTTCTGTGTTTGAATGTTCT 480 |
| DB | 421 TTGCTAAGCCATTTGATCAAGAGGTTATCAAGGGTCTTCTGTGTTTGAATGTTCT 480 |
| QY | 481 ATGCTGCTGCAATGTCACAGATCTCTCAAAACCAAGTTGTTTACAACATGTCAGAGCT 540 |
| DB | 481 ATGCTGCTGCAATGTCACAGATCTCTCAAAACCAAGTTGTTTACAACATGTCAGAGCT 540 |
| QY | 541 GGTGTTATCCATTTGGCTTGAAGCTTGGCTTGAATGGGCTTAAGTACAACATCAGAGTT 600 |
| DB | 541 GGTGTTATCCATTTGGCTTGAAGCTTGGCTTGAATGGGCTTAAGTACAACATCAGAGTT 600 |
| QY | 601 AATTCTTTAAACCCAGTTACATCTACGGTCTTTGACCAAGATGTTATCAATGTTAAC 660 |
| DB | 601 AATTCTTTAAACCCAGTTACATCTACGGTCTTTGACCAAGATGTTATCAATGTTAAC 660 |
| QY | 661 GAAGAAATGTAACAACAGATGATCTGTGATTCACCAACAAAGATGTCGAACCAAG 720 |
| DB | 661 GAAGAAATGTAACAACAGATGATCTGTGATTCACCAACAAAGATGTCGAACCAAG 720 |
| QY | 721 GAATACATTTGTCGCTTTTGTACTGCTTCTGATCTGCTGCTCATACACTACTGCT 780 |
| DB | 721 GAATACATTTGTCGCTTTTGTACTGCTTCTGATCTGCTGCTCATACACTACTGCT 780 |
| QY | 781 GCCAGCTTACTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 816 |
| DB | 781 GCCAGCTTACTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 816 |

| | |
|-----------------------|---|
| ORIGIN | |
| Query Match | 40.8%; Score 332.8; DB 8; Length 1166; |
| Best Local Similarity | 65.0%; Pred. No. 1.7e-73; |
| Matches | 528; Conservative 0; Mismatches 272; Indels 12; Gaps 2; |
| QY | 14 TTCCAACTTTAGATTCGATGCGCACTTAACCATTTGTCACAGGTCCTGTGCTTTAG 73 |
| DB | 278 TTCCAACTTTAGATTCGATGCGCACTTAACCATTTGTCACAGGTCCTGTGCTTTAG 337 |
| QY | 74 CTGAAGCTTTATCAAGGTTTGTGGCTTACGGTTCTGACATGCTTGTGCTGATATCG 133 |
| DB | 338 CCGCTGTTGTTCCAGAACCTTTGTAACAAAGGCTGATGATGCTTGTGATGATGATGA 397 |
| QY | 134 ACCAGAAAGACGCTGTCGCAAAACAAGCCGATACCAAAATACGCTAATGGAATGGA 193 |
| DB | 398 ACTTGGAAGAAACCAACAAGCCGCAAGAGATGTTTACATGGGGTGAAGAAACAATGA 457 |
| QY | 194 AGTTGAAAGA-----AGTTCCAAAGATGGGTTCAATAGCTGTGATATTTCTGATT 244 |
| DB | 458 AAGAAACACAGAACTCCAAATTTGTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGATG 517 |
| QY | 245 CTGATACCGTTTACAAAGGTTTGTCTCAAGTTGCTCAAGATTTTGTGTAAGTTGCCATTGC 304 |
| DB | 518 CTGAAGCTGTTGAATTAATGACATTCAAAGCCATACGAACCAACGCTTAAGTTGCCAGTG 577 |
| QY | 305 ACTGGTTTAAACAGAGCTGTTACTGTGAATACTTCCCATGTGAAGATTACCAAGCCACA 364 |
| DB | 578 TTTTGATTAACACAGCTGTTATGCTGAATAATTTCCCTGCTGAAGAAATACCAAGCCACA 637 |
| QY | 365 ACGCTGGAAGATGTTGAAGGTTAACTTGTGGGTTCTTTGATGTTTCTCAAGCCTTTG 424 |

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.

1 (bases 1 to 1166)

Murray, J.S., Wong, M.L., Miyada, C.G., Switchenko, A.C., Goodman, T.C. and Wong, B. Isolation, characterization and expression of the gene that encodes D-arabinol dehydrogenase in Candida tropicalis

Gene 155 (1), 123-128 (1995)

95212917

7696655

2 (bases 1 to 1166)

Wong, B. Submission

Submitted (10-AUG-1993) Brian Wong, Internal Medicine, University of Cincinnati College of Medicine, 231 Bethesda Avenue, Cincinnati, OH 45267-0560, USA

FEATURES

source

1..1166

/organism="Candida tropicalis"

/mol_type="genomic DNA"

/strain="ATCC 750"

/db_xref="taxon:5482"

/clone="pCT1A"

/gene_1db="genomic DNA library in pAB107 from V.F. Kalb and J.C. Loper"

241..1089

/gene="ARD"

241..1089

/gene="ARD"

/function="catalyzes D-arabinol + NAD <=> D-ribulose + NADH"

/codon_start=1

/evidence=experimental

/product="D-arabinol dehydrogenase"

/protein_id="AA6635.1"

/db_xref="GI:392786"

/translation="MDSSSYWSYDNIYPSRLDKLYITIGSGSLSAVSRALLAGADIALIDMNLERTQAAADVYLQWGEQMKHESPIQVSNMNCIGDAEVELTFLA INEHGKVAAYLINTAGYAENFPAEYPAKAEINIKVNGISPYVQAFARPLQNN MTGSIILIGMSGTIVNDPOCKYNNMSKGVILHSLACEMAKYNI RAVTSLSPGYI LTPITRNVISHTEMKTEWESKIPMKMAEKERVGSILVLAISDASASYTTGHNLYVD GYEWCW"

gene

CDS

```

Db      638 ATCCGAAATATTCATGAAGTCATGTTAGATCTTTACGCTTCAACAGCTTTG 697
Qy      425 CTAAAGCATATGATCAAGAGATATCAAGGTCCTTCTGTTTGAATGTTCTATGT 484
Db      698 CCAAGCCTTGAATCAAAA---TAACTAGCTGGCTCTATATCTTGAATGTTCTATGT 754
Qy      485 CTGGTCCATGTCACAGATCTCTCAAAACCAAGTTGTCTACACATGTCACAGGCTGTG 544
Db      755 CTGGTACCATTTGTCAACGATCCCAACCAATGATGACCAATGTCACAGGCTGTG 814
Qy      545 TTATCATTTGGCTAGACTTGGCTGTGATGAGGCTAAGTCAACATCAAGTTAAT 604
Db      815 TCATTCATCTGGCCAGATTTTGAAGATGGAATGGCCAAATACACATCAAGTCAATA 874
Qy      605 CTTTAAACCCAGGTTATCACTACGCTCTTGAACCAAGATGTTATCAATGTTAAAGAA 664
Db      875 CTTTCCCGCAGGTTATATTTTGAATCTCCATTCACAAAGAAATTTATAGTGTCACTG 934
Qy      665 AATTGTACACAGATGATCTGTGTATCCCAACCAAGAAATGTCGAAACCAAGAAAT 724
Db      935 AATATGAACAGAAATGGATCCAGATTCAGATGAGAAATGGCCGACCAAAAGAT 994
Qy      725 ACATGGGCTGTTTGTATCTGCTTCTGATCTGCTGCTTCACTACTGCTGCTCA 784
Db      995 TTGCGGTTCTATTTTGTACTGCTGCTTCTGATCTGCTCTTCACTACTGCTCA 1054
Qy      785 GCTTACGTTGATGATGCTGCTTCACTCTTGG 816
Db      1055 ATTGCTAGTGAATGCTGTTACGAATGTTG 1086

```

RESULT 3
 YSAARDH 1614 bp DNA linear PLN 21-JUN-1995
 LOCUS Candida albicans D-arabinitol dehydrogenase (ARDH) gene, complete
 DEFINITION cds.

ACCESSION L16227
 VERSION L16227.1 GI:295567
 KEYWORDS Ardh gene; D-arabinitol dehydrogenase.
 SOURCE Candida albicans
 ORGANISM Candida albicans

REFERENCE Wong, P., Leeson, S., Grindle, S., Magee, B., Brooks, E. and Magee, P.T.
 1 (bases 1 to 1614)
 TITLE D-arabitol metabolism in Candida albicans: studies of the biosynthetic pathway and the gene that encodes NAD-dependent D-arabitol dehydrogenase

JOURNAL J. Bacteriol. 175 (19), 6314-6320 (1993)
 MEDLINE 94012494
 PUBMED 8407803

REFERENCE Wong, P., Leeson, S., Grindle, S., Magee, B., Brooks, E. and Magee, P.T.
 2 (sites)
 TITLE D-arabitol metabolism in Candida albicans: construction and analysis of mutants lacking D-arabitol dehydrogenase

JOURNAL J. Bacteriol. 177 (11), 2971-2976 (1995)
 MEDLINE 95286472
 PUBMED 7768790

REFERENCE Wong, B.
 3 (bases 1 to 1614)
 TITLE Direct Submission

JOURNAL Submitted (18-MAY-1993) Brian Wong, Department of Internal Medicine, University of Cincinnati College of Medicine, OH 45267-0560, USA
 COMMENT Original source text: Candida albicans (strain WO1) DNA.

FEATURES
 source
 1..1614
 location/Qualifiers

/organism="Candida albicans"
 /mol_type="genomic DNA"
 /strain="WO1"
 /db_xref="taxon:5476"
 /map="6; segment O"

```

gene      /cbase_1id="pEMBLv23 from P.T. Magee"
          /gene..1280
          /gene="ArDH"
          /size..1280
          /gene="ArDH"
          /function="catalyzes D-arabinitol + NAD = D-ribose + NADH"
          /codon_start=1
          /evidence=experimental
          /transl_table=12
          /product="D-arabinitol dehydrogenase"
          /protein_id="AAC37430.1"
          /db_xref="GI:295568"
          /translation="MDSAYMSYDINIVPSFRIDGKLVITGGSGGLAAVSRALLAKGA
          DVALVDNMLERTQOARADVLOMGEOMKGTESPIGVASMSCNIGDAVDLFEKAI
          NEHGKISVIVNAGVKNPFAEYKPMKMLNKKVGLSPYQSAFAPLIONNM
          TGSILIGSMGCTIVNDPOCMYNTMSAGYTHLARSLACSMKATNIVNTLSPEYIL
          TPLRNIVSGHTEKTEWESKIPKRMHAPKPFVGSILYLAESASSTTGHNLVVDG
          GYECW"

```

Query Match 37.5%; Score 305.6; DB 8; Length 1614;
 Best Local Similarity 62.9%; Pred. No. 1.3e-66;
 Matches 511; Conservative 0; Mismatches 289; Indels 12; Gaps 2;

```

Qy      14 TTCCAACTTTAGATTCGATGCGCACTTAACATTTGTCACAGTGCCTGTGGTTTAA 73
Db      469 TCCCAAGTTCGCTTGGATGGAAGAACTAGCATATTAACCGGTGCTCTGTGGTTTGG 528
Qy      74 CTGAAGCTTTATATCAAGGCTTGTGGCTTCAAGGCTTCAAGTTCCTTCTTGAATTC 133
Db      529 CTGCGGTGTATCAAGAGCTTTATTAAGCCAAAGGCGCATGTTGATTAATGATGA 588
Qy      134 ACCAAGAAAGACGCTGCTCCAAACAGCCGAATCCAAATAGCTACTGAAGATGA 193
Db      589 ACTGGAAGAAACACACACAGCTGCTAGAGAGCTTTACAAATGGGCGCAAGACCAATGA 648
Qy      194 AGTGAAGAAAGTTCCAA-----AGATGGGCTTATGCTGTGATATTTCTGATTT 244
Db      649 AAGGAAATATAGATTCACCAATCGGTGAGTGAAGTCTGTGATATTAATTTGGCCATG 708
Qy      245 CTGATACGTTTCAACAGGCTTGTGCTCAAGTGTCTGAAGATTTTGTGAATGCTCCATTGC 304
Db      709 CTGAAGCTGTGACTTTCATTTCAAGGCTATCAACGACCAACCAACGCAAAATTCGAATG 768
Qy      305 ACTGTGTTAACAAGCTGTTACTGTGAATACTTCCCATGTGAAGATTAACCAAGCAAGA 364
Db      769 TCTTGTCAACACGCGGTTTACGCTGAATACTTCCAGTGAAGATTAACCAAGCAAGA 828
Qy      365 ACGCTGAGAAAGATGATGAGGTTTAACTTGTGGGCTCTTGTATGTTTCTCAAGCTTTG 424
Db      829 ACGCTGAGAAACCTTATATGAATTAACGGGTTGGGCTCATTTACGTTTCCCAAGCTTTG 888
Qy      425 CTAAAGCATATGATCAAGAGATATCAAGGCTGCTTCTGTTTGAATGTTCTATGT 484
Db      889 CTAAAGCATATTAACAAA---CAACATGACCGGATGATCATTTGATCGGGTCAATGT 945
Qy      485 CTGGTCCATGTCACAGATCTCTCAAAACCAAGTTGTCTACACATGTCACAGGCTGTG 544
Db      946 CCGGTACATCTGTAAAGACCAACCAACCAATCATGTATCAACATGTCACAAACCGGTG 1005
Qy      545 TTATCATTTGGCTAAGACTTGGCTGTGATGAGGCTAAGTACAACTCAAGATTAAT 604
Db      1006 TCATTCATTTAGCCAGATCATTTGCTGTGATATGAGGCTTAATCAATATCAAGATCAACA 1065
Qy      605 CTTTAAACCCAGGTTATCACTACGCTCTTGAACCAAGATGTTATCAATGTTAAAGAA 664
Db      1066 CATTTGCGCCAGGTTATCACTTAAACCATTTGAACAAAGAACTTATTAAGTGAACACCG 1125
Qy      665 AATTGTACACAGATGATCTGTGTATCCCAACAAAGAAATGTCGAACCAAGAAAT 724
Db      1126 AATGAAAGACAGATGGAATCAAAAGATCCCAATGAAGAAATGGAACCAAAAGAT 1185

```



```

Oy 5 CTGACTACATCCAACTTTAGATTGATGCGCACTTAACATGTGACAGTGCCTG 64
Db 207 CTAAAGTGTGTTCCCACTTCAATTCAGATGGAGAGATGGCTATTTATCCGAGGTTCTG 266
Oy 65 GTGGTTTAGCTAGACCTTTATCAAGGGTTTGTGGCTTACAGGTTTGTGACATTTGCTTGC 124
Db 267 GTGGTTTAGCTAGACCTTTATCAAGGGTTTGTGGCTTACAGGTTTGTGACATTTGCTTGC 326
Oy 125 TTGATATCGACCAAGAAAAGACTGTGCTCCAAACAAGCCGATATACCAAAATACGCTACTG 184
Db 327 TTGATATCGAACTTGGAAAAGAACCAAGTCGCTGCCAAAGAAATTTGGCTGGGGGAAG 386
Oy 165 AAGATTTGAAGTTGAAAAGAGTTCCAAAGATGGGTTCAATG-----CCTGTGATA 235
Db 387 AGACGTTGAAGGGTGAACACGCTTCAGCCATCGGCCAAGTTTCCGCTGGTCTTGCAACA 446
Oy 236 TTTGATTTCTGATACCGTTTCAAGAGTGTGCTCAAGTTGCTAAGATTTTGGTAAGT 295
Db 447 TTGGGGAATGCTAGGCACTAGACGCTACTTTCAAGCTCCATCAACGAACCAACGCGCAAGA 506
Oy 296 TGCATTCGACTTGTGTTAACACAGCTGTGTTACTGTGAAAACCTCCATGTGAAGATTAC 355
Db 507 TCGCTGACTGTGTGATTAACACGCTGGATCTGTGAAAACCTCCGCGCAAGAGTACC 566
Oy 356 CAGCCAAAGACGCTGGAAGAGTGTGAAGTTAACTTTGGGTTCTTTGATGTTTCTC 415
Db 567 CGGCTACTTAACGCTTAAGCATGAAAGTGAACGGTTTGGCTCATTTCTACGTTTGC 626
Oy 416 AAGCTTTGCTAAGCATTTGATCAAAAGATATCAAGGGTCTCTGTGTTTGTGATG 475
Db 627 AATCGTTGCTAGACATTTGATTCAGAA---CACTTGAAGGCTCTATCATCTTGATG 683
Oy 476 GTTCTATGTCTGTGTCATTTGTCACAGATCTCAAAACCAAGTTGCTCAACATGTCCA 535
Db 684 GCTCATATGTCTGGAACATTTGTCAACGACCAACCCCAAGTATGTACACATGTCCA 743
Oy 536 AGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGCTAAGTCAATCA 595
Db 744 AGGCTGTGTATTCATTTGGCTAAGACTTTGGCTTGTGAATGGCTAAGTCAATCA 803
Oy 596 GAGTTAATCTTTAAACCAAGTTACATCTACGGCTTTGACCAAGATGTATCAATG 655
Db 804 GAGTTAATCTTTAAACCAAGTTACATCTACGGCTTTGACCAAGATGTATCAATG 863
Oy 656 GTAAAGAAATTTGACAAAGATGATCTGTGATCCCAACAAGAAATGTCCGAAC 715
Db 864 GCCACAGAGATGAAAGAGAGCTGTGAATCCCAAGATCCCAAGAAATGTCCGAAC 923
Oy 716 CAAGGAATTAATGTGTCTGTTTGTACTTGTCTTGTGAATCTGTCTCTTCAACTCA 775
Db 924 CAAGGAATTAATGTGTCTGTTTGTACTTGTCTTGTGAATCTGTCTCTTCAACTCA 983
Oy 776 CTGGTCCAGCTTACTGTGTGATGTGTGTT 805
Db 984 CGGCGCACAATTTGTGTGACGGAAGAT 1013

RESULT 6
AR031556 696 bp DNA linear PAT 29-SEP-1999
LOCUS AR031556
DEFINITION Sequence 6 from patent US 5866382.
ACCESSION AR031556
VERSION AR031556.1 GI:5945845
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 696)
AUTHORS
Hallborn,J., Penttila,M., Ojamo,H., Walfridsson,M., Airaksinen,U.,
Keränen,S. and Hahn-Hagerdal,B.
TITLE
Xylose utilization by recombinant yeasts
PATENT: US 5866382-A 6 02-FEB-1999;
FEATURES
Location/Qualifiers

```

```

source 1. 696
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 29.2%; Score 238.2; DB 6; Length 696;
Best Local Similarity 64.4%; Pred. No. 1.7e-49;
Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

Oy 227 CCTGTGATATTTCTGATTCGATACCGTTTCAAGAGTGTGTTGCTCAAGTGTGAAGATT 286
Db 107 CTTGCAACATTGGGAGATGCTGAGCGATGAGACGCTACTTTCAAGCTTCATCAACGAACACC 166
Oy 287 TTGTGAATGTGCACTTGTGTTAAACAGCTGTGTTACTGTGAAAACCTTCCATGTG 346
Db 167 ACGGCAAGATGCTGACTGTTGATTAACCGCTGGAATACGTGAAAACCTTCCGCGC 226
Oy 347 AAGATTACCCAGCCCAAGACGCTGGAAGATGTGAAGTTAACTTTGGTGTCTTTGT 406
Db 227 AAACGTACCCGCTACTAAGCGCTGAACCATCATGMAAGTGAACGTTTGGCTCATTTCT 286
Oy 407 ATGTTTCTAAGCCTTGTGCTAAGCATGATCAAAAGATATCAAGGGTGTCTGTG 466
Db 287 ACGTTTCCGAATCGTTCGCTAGACCATGATTCAGAA---CACTTGAAGGCTCTATCA 343
Oy 467 TTTTGATTTGTTCTATGTCTGTGTCATTTGTCACAGATCTCAAAACCAAGTTGCTTACA 526
Db 344 TCTTGATTTGGCTCAATGCTGTGAACATTTGTCAACGACCCCAACCAAGTATGTACA 403
Oy 527 ACATGTCCAAGGCTGTGTATTCATTTGCTAAGACTTTGGCTTGTGAATGGCTAAGT 586
Db 404 ACATGTCCAAGGCTGTGAATTCACATTTGCTGATCGTGTGCGCTGCAATGGGCCAAGT 463
Oy 587 ACAACATGAGATTAATCTTTAAACCAAGTTACATCTAGCGTCTTGAACCAAGATG 646
Db 464 ACAACATGAGATTAATCTTTAAACCAAGTTACATCTAGCGTCTTGAACCAAGATG 523
Oy 647 TTATCAATGTGTAAAGAAATTTGTAACAAGATGATCTGTGATCCCAACAAGAAAGAA 706
Db 524 TGATTTTCTGGCCACACAGAGATGAAAGAGCTGGGATCCAAAGATCCCAAGAAAGAA 583
Oy 707 TGTCCGAACCAAGAAATTAATGTGTCTGTTTGTACTTGTCTTGTGAATCTGTGCTT 766
Db 584 TGGCCGAACCAAGAAATTAATGTGTCTGTTTGTACTTGTGAATCTGTGCTTGT 643
Oy 767 CATACACTGTGTGCGCACTTACTGTGTGATGTGTGTT 805
Db 644 CTAACACTGAGGCGCCACAAATTTGTGTGACGGAAGAT 682

RESULT 7
AR345028 696 bp mRNA linear PAT 17-AUG-2003
LOCUS AR345028
DEFINITION Sequence 6 from patent US 6582944.
ACCESSION AR345028
VERSION AR345028.1 GI:33741148
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 696)
AUTHORS
Hallborn,J., Penttila,M., Ojamo,H., Walfridsson,M., Airaksinen,U.,
Keränen,S. and Hahn-Hagerdal,B.
TITLE
Production of ethanol from xylose
PATENT: US 6582944-A 6 24-JUN-2003;
FEATURES
1. 696
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 29.2%; Score 238.2; DB 6; Length 696;
Best Local Similarity 64.4%; Pred. No. 1.7e-49;

```

Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 227 CCCTGATATTTGATTCGATACCGTTCACAGGTTTGTCCAGTTCTCAAGGTT 286
DB 107 CCTCAACATTGGGAGTGGAGTAGAGCGTACTTTCAGCTCCATCAACGACACC 166
QY 287 TTGGTAAGTCCCATTCGATCTGGTTAACAGAGCTGTCTTCTGAAAACCTTCCATGTG 346
DB 167 ACGGCAAGATGGCTGACCTGTGTTAATCAACCGGTGATACGTGAAAACCTTCCGCG 226
QY 347 AAGATTACCCAGCAAGACGCTGAGAAAGATGTGAGGTTAACTTGTGGCTTCTTGT 406
DB 227 AAGAGTACCGGCTTACTAACCTGAAACATCAAGAAAGTGAACGGTTTGGCTCATCT 286
QY 407 ATGTTCTCAGCCCTTGTCTAAGCATGATCAAGAGATCAAGAGGCTCTCTGTG 466
DB 287 ACGTTTGCATTCGTTCCGTAGACCATTTGATCCAGAA---CACTTGAGAGGCTTATCA 343
QY 467 TTTTGAATGTTCTAATGTCTGGTCCATTTGCAAGATCTCAAAACCAAGTTGTCTACA 526
DB 344 TCTTGAATGGCTCAATGTCTGGAACAATTTGCAACGACCAACCCCAATGATGTACA 403
QY 527 AATGTCCAGAGGCTGTGTATTCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGT 586
DB 404 AATGTCCAGAGGCTGTGTATTCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGT 463
QY 587 ACAATCAGAGTTAATCTTTAAACCAAGTTACATCAAGGCTCTTGAACCAAGATG 646
DB 464 ACAATCAGAGTTAATCTTTAAACCAAGTTAATTTGACTCTTTAAACCAAGATG 523
QY 647 TTATCAATGTGAACGAAATTTGTAACAAGATGATCTGTGTAATCCCAACCAAGAA 706
DB 524 TGATTTTGGCCACACAGAGATGAAGAGAGCTGGGAATCCAAAGATCCCAATGAAGAA 583
QY 707 TGTCCGAACCAAGAAATTAATGATGCTGTTTGTACTGTCTTCTGAATGTGCTT 766
DB 584 TGGCCGAACCAAGAAATTTGTTGGGCTCATCTTAACTTTGGCAAGCAAGACTGCTT 643
QY 767 CATACACTAGTGTGCAAGCTTACTGTTGATGTGTT 805
DB 644 CTTACACTAGTGTGCAAGCTTACTGTTGATGTGTT 682

RESULT 8
AR544898 846 bp DNA linear PAT 08-OCT-2004
LOCUS AR544898 Sequence 29 from patent US 6747137.
DEFINITION AR544898
ACCESSION AR544898
VERSION AR544898.1 GI:53938073
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 846)
Weinstein, K.G. and Bush, D.
Nucleic acid sequences relating to Candida albicans for diagnostics
and therapeutics
JOURNAL Patent: US 6747137-A 29 08-JUN-2004;
FEATURES
location/Qualifiers
source
1..846
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 26.8%; Score 218.6; DB 6; Length 846;
Best Local Similarity 62.0%; Pred. No. 1.5e-44;
Matches 385; Conservative 0; Mismatches 224; Indels 12; Gaps 2;

QY 14 TTCGAATTTGATTCGATTCAGACTTAAACATTTGTCAAGAGCTGTGGGTTAG 73
DB 218 TCCCAAGCTTCGATTCGATTCAGACTTAAACATTTGTCAAGAGCTGTGGGTTAG 277
QY 74 CTGAAGCTTTAATCAAGGCTTGTGGCTTACGCTTCAAGATGTGCTTGTGATATGC 133

DB 278 CTGCGGTGATCAAGAGCTTTATTTAGCCAAAGGTGCGATGTGCAATGTCGATATGA 337
QY 134 ACCAAGAAAAGACTGCTGCCAACAAGCCGAATACCAAAATACGTAAGAAATTTGA 193
DB 338 ACTTGGAAAAGAACCAACAAAGCTGTAGACGCTTTACAAATGGGGGAAGACAAATGA 397
QY 194 AGTTGAAAAGAGTTCCAA-----AGATGGTTCAATATGCTGTGATTTTGTGATT 244
DB 398 AAGGTAAATATGCAATCAACATCGTCAAGTGAAGTGTCTGTGATTAATATGGCCATG 457
QY 245 CTGATACCTTTCACAAAGTGTGTCTCAAGTGTGATGATTTTGTGAATGGCATTCG 304
DB 458 CTGAAGCTGTGACTTTGACATTTCAAGCCATCAACGAAACCAACGCAAAATCTCAAGTG 517
QY 305 ACTTGTAAACACAGCTGTGTTAATCTGTGAAAACCTTCCATGTGAAGATTACCAACCAAGA 364
DB 518 TCTTGTTCACACATGCGCGGTTTACGCTGAAAACCTTCCAGCTGAAGATTAACCAACCAAGA 577
QY 365 ACGCTGAAGAGATGTGAAGTTAACTTGTGGCTTTTGTATGTTTCTCAAGCTTTG 424
DB 578 ACGCTGAAGAGATGTGAAGTTAACTTGTGGCTTTTGTATGTTTCTCAAGCTTTG 637
QY 425 CTGAAGCTGTGATCAAGAGATTTCAAGGCTGCTTGTGTTGATTTGATTTGATTTGAT 484
DB 638 CTGAAGCTGTGATTTTCAAAA---CAACATGACCGATGATATTTGATGCGGTCAATGT 694
QY 485 CTGGTGCATTTGTCAAGATCTCTCAAAACCAAGTTGTCTCAACATGTCCAAAGCTGTG 544
DB 695 CCGTACCATGTGTACAGACCCCAACCAACCAATGATGATCAATGTCCAAAGCCGCTG 754
QY 545 TTATCAATTTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGTCAATCAATCAAGTTAAT 604
DB 755 TCATTCATTTAGCCAGATCATTTGGCTGTGATGGGCTTAATATCAATATGAGATCAACA 814
QY 605 CTTTAAACCCAGTTTACATCT 625
DB 815 CATTGTGCGCGGACACATCT 835

RESULT 9
CR382121_09/c
WPCOMMENT
Sequence split into 11 fragments LOCUS CR382121 Accession CR382121
Fragment Name Begin End
CR382121_00 1 110000
CR382121_01 100001 210000
CR382121_02 200001 310000
CR382121_03 300001 410000
CR382121_04 400001 510000
CR382121_05 500001 610000
CR382121_06 600001 710000
CR382121_07 700001 810000
CR382121_08 800001 910000
CR382121_09 900001 1010000
CR382121_10 1000001 1062590
Continuation (10 of 11) of CR382121 from base 900001 (CR382121 Kluveromyces fragilis strain)

Query Match 19.1%; Score 155.8; DB 8; Length 110000;
Best Local Similarity 53.4%; Pred. No. 1.1e-28;
Matches 473; Conservative 0; Mismatches 337; Indels 75; Gaps 4;

QY 1 ATGACTGACTACATTTCAACTTTTATGATTCAGTGGCACTTAACCATTTGTCAAGAGTCC 60
DB 84309 ATTACTGATGTTTGTGCTCTTTTATGATTTGATGACAGTAAACATGTCACGTGCT 84250
QY 61 TGTGTGTTTATGCTGAAGCTTTTAAATCAAGGCTTTTGTGGCTTACGCTTGTGACATTTGCT 120
DB 84249 GCCGTGTGTTGGCTTCAACATTTGTCAACAGGCAATGTGTCAACGGGATTCGAACATGCA 84190
QY 121 TTGCTGATATTCAGCAAGAAAGATCTGTGCAAGAGCAATATACCAAAATATAGCT 180
DB 84189 TTGATGATCTTTCTTGTGCTTATGAGAAATGTTAAATCTGAATGTGAATGAACTTTC--- 84133

CR382133_03 300001 410000
CR382133_04 400001 510000
CR382133_05 500001 610000
CR382133_06 600001 710000
CR382133_07 700001 810000
CR382133_08 800001 910000
CR382133_09 900001 1010000
CR382133_10 1000001 1110000
CR382133_11 1100001 1210000
CR382133_12 1200001 1249565
Continuation (12 of 13) of CR382133 from base 1100001 (CR382133 Debaryomyces hansenii ch

Query Match 11.3% Score 92.6; DB 8; Length 110000;
Best Local Similarity 52.5%; Pred. No. 1.1e-12;
Matches 309; Conservative 0; Mismatches 259; Indels 21; Gaps 4;

QY 221 CATATGCTGTGATTTCTGATTCGTATACCGTTCAACAAGTGTTCCTCAAGTTGCTA 280
DB 99071 CTTATTAATGCTGCTGATTAATTTTGAAGAAGTAGAAGACCTGTAATCAATAGTGA 99012
QY 281 AGAATTTTGTAGTGCCTTCACTTGTGTTAACAAGCTGTGTTACTCTGTAATACTTCC 340
DB 99011 AGAATTTTGAACAATGATTAATTCGTGCTAATGCTGGGGTGTGCGAATCTGCG 98952
QY 341 CATGGAAGATTACCCAGCAAGACGCTGAAGAAGATGTAAGGTTAATCTTGTGGT 400
DB 98951 CAGTAAGGAGCTTCTTATGCTTACCAATGCTATGTAAGTTAATCTTACGCGC 98892
QY 401 CTTTGTATGTTCTTCAGCCTTTCCTTCAAGCCTTATGATCAAGATCAAGGCTCTT 460
DB 98891 TTTACTATGCTCTCTTCCCAAGCTTATGACCAATTTTCGAAGAATGGGAAGAT 98838
QY 461 CTGTTGTTTGTAGTGTCTATGCTGTGCTGCAATGCTCAAGATCTCTCAAAACCAAGTTG 520
DB 98837 CATTAATTAATTAATCTCTTCCCAAGCTTATGACCAATGCTCTTATGCTCAAGCTG 98778
QY 521 TCTACAACATGTCGAAGGCTGTGTTATCCATTTGGCTTAAGCTTGTGCTGTGATGAG 580
DB 98777 CTTAATTAAGCAATTAAGCAAGCTCTTATCCATTTGGGAATCACTAGCAATGAGTGG 98718
QY 581 CTAAGTACAACTCAGAGTTAATCTTTAAACCAAGTTCACTCAAGCTCTTGTGACCA 640
DB 98717 TTAAGT---TTGCAAGATTAATTAATCACTCAAGTTAAT---TTTGACCG 98670
QY 641 AGAATGTTATCAATGTGAAGAAATGTTAACAAGATGATCTCTGTATCCCAAC 700
DB 98669 AGATTGCGAACAATGTTCCCTTGTGATGTAAGTCAAAAGTGCGCTTTAATTTCAATGG 98610
QY 701 AAGAATGTCGAACCAAGAATAATCATGTTGCTGTTTGTACTTGTCTGAATCTG 760
DB 98609 GGAAGAGGCTGCGCAAGAATTAATGCGAGCATTTTGTATTTTGTGATCTGAC---G 98553
QY 761 CTGCTTCAATACACTGCTGTCAGCTTAACTGTTGATGATGATGATGATGATGATG 809
DB 98552 CATCTACATTAATCACTACTGCTGCAAGATTAATTAATGATGCGGATATAC 98504

RESULT 12
CR382137_06
WPCOMMENT
Sequence split into 21 fragments LOCUS CR382137 Accession CR382137

Fragment Name Begin End
CR382137_00 1 110000
CR382137_01 100001 210000
CR382137_02 200001 310000
CR382137_03 300001 410000
CR382137_04 400001 510000
CR382137_05 500001 610000
CR382137_06 600001 710000
CR382137_07 700001 810000
CR382137_08 800001 910000
CR382137_09 900001 1010000
CR382137_10 1000001 1110000

CR382137_11 1100001 1210000
CR382137_12 1200001 1310000
CR382137_13 1300001 1410000
CR382137_14 1400001 1510000
CR382137_15 1500001 1610000
CR382137_16 1600001 1710000
CR382137_17 1700001 1810000
CR382137_18 1800001 1910000
CR382137_19 1900001 2010000
CR382137_20 2000001 2037969
Continuation (7 of 21) of CR382137 from base 600001 (CR382137 Debaryomyces hansenii chro

Query Match 11.2% Score 91.4; DB 8; Length 110000;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 331; Conservative 0; Mismatches 276; Indels 24; Gaps 5;

QY 182 CTGAAGAATTGAAGTTGAAGAAGTTCCAAAGATGGGTTATATAGCTGTGATTTCTG 241
DB 92817 CTGAATCTTAACCAACCACTTACGAGTGAATTCGAAGCATACAGTGTATATTTTCG 92876
QY 242 ATTCTAACCGTTCAACAAGGTTTGTCTCAAGTTGCTAAGATTTTGTAGTTGCCAT 301
DB 92877 ACCCAAGAAGTGTGAAGAAGGTATGGAACAATTTGAAGAATTTGCGCATTTGACG 92936
QY 302 TGCACTGTGTTACACAGCTGTGTTACTGTGAAAATCTCCATG---TGAAGATTACCCAG 358
DB 92937 TTTTGTGCTAATGCTGTGTTGCTGATGACCGAGGCAAGATTTGAAGCGGATGGT 92996
QY 359 CCAAGAAGCTGAAGAAGTGTGAAGTTAATCTTGTGGTCTTGTATGTTCTCAAG 418
DB 92997 ACGATGCGTGAAGAAGTGTGATTTGATTTGACCTTGAATGAGTTACTACGCGCTAAG 93056
QY 419 CTTTGTACGCCATTTGATCAAGAAGTATCAAGGCTCTGTTGTTTGTATGTTGTT 478
DB 93057 CTGTCGGAAGATTTTCAAGAAGCACGAA-----AGGGTCTCTGTGTTAACAAGCAT 93110
QY 479 CTATGCTGTGCTCAATGCTCAAGATCTCTCAAAACCAAGTTGTTCAACAATGCTCAAG 538
DB 93111 CAATGCTGCGCAATGTTGATGATGTTCCCAAAAGCAAGCTTCAPACCAAGCTCTGTAAG 93170
QY 539 CTGCTGTTATCAATTTGGCTAAGCTTGGCTTGTGATGAGGCTTAAGTCAACATCAAG 598
DB 93171 CGGGTATTTCACTGAGTGAATGCTTACCTTACATGCAATGAGGCTCCATTCGCT---AGAG 93227
QY 599 TTAATCTTTAAACCAAGTTTACATCTCAAGCTCTTGTGACCAAGAATGTTAATGATG 658
DB 93228 TTAACACATTTCTCCAGATTACAT-----TGTACAGAGATTTCCGATTTTGTTC 93278
QY 659 ACGAAGAATGTTACAACAGATGATCTGTGATTCCTCAACAAGAATGTCGGAACGA 718
DB 93279 CTGCCGAAGTCAAGGCTAAATGTTGCAATTTGATCCATTGGAAGAAGACACTTGCTC 93338
QY 719 AGGAATACATTTGCTGCTGTTTGTACTGTTTGTGATGCTGCTTCAATGCTCATACACTG 778
DB 93339 AAGAGTTGTTGTTGCTTACTTAATTTTGTGATCCGA---TCTTCTACTTACACAGACTG 93395
QY 779 GTGCCAGCTTACTGTTGATGTTGTTTTCAC 809
DB 93396 GCTCGGACTTAATGTTGATGTTGTTGTTACTC 93426

RESULT 13
AF002134
LOCUS
DEFINITION
Candida albicans Sou2p (SOU2), Sou1p (SOU1) and Vma8p (VMA8) genes,
complete cds.
ACCESSION
AF002134
VERSION
AF002134.1 GI:2183241
KEYWORDS
SOURCE
ORGANISM

Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE 1 (bases 1 to 4293)
AUTHORS Jandon, G., Sherman, F. and Ruschenko, E.
TITLE Monosomy of a specific chromosome determines L-sorbose utilization:
a novel regulatory mechanism in *Candida albicans*
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (9), 5150-5155 (1998)
MEDLINE 9826783
PUBMED 9560244
REFERENCE 2 (bases 1 to 4293)
AUTHORS Jandon, G., Ruschenko, E. and Sherman, F.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-1997) Department of Biochemistry and Biophysics,
University of Rochester, 601 Elmwood Avenue P.O. Box 712,
Rochester, NY 14642, USA
FEATURES
SOURCE Location/Qualifiers
1. .4293
/organism="Candida albicans"
/mol_type="genomic DNA"
/strain="Sor17"
/db_xref="taxon:5476"
/chromosome="4"
356. .1199
/gene="SOU2"
356. .1198
/gene="SOU2"
/feature="short-chain alcohol dehydrogenase homolog"
/codon_start=1
/transl_table=12
/product="Soulp"
/protein_id="AAC24462.1"
/db_xref="GI:2183242"
/translation="MSKETTSYTNKAGLPPTKATIPDNLDAFSLKGVKASVTGSS
GGIGMAVAEGYAGADVAIMYNHSPADDAEYLTGYGVSKAYKCNVDPDVEKY
VKQJESFGTIDIFVANAGVAMTEGPEKFDVGRKMNKVNVDLNSVYCAHVGP
RKKSGSGFIPFASMSASIVNVPOLAAVNAKAGVGLSKSFVEMAPFARVNSVSP
YIATHLEFPADPDVSKWMLQPLGREKREKRELVGAVLYLASDAATTTGADLAVDGS
VTYV"
2063. .2909
/gene="SOU1"
2063. .2908
/gene="SOU1"
/feature="short-chain alcohol dehydrogenase homolog"
/codon_start=1
/transl_table=12
/product="Soulp"
/protein_id="AAC24463.1"
/db_xref="GI:2183243"
/translation="MSREIISFTNPALGPLPTKAPOLPSNVLDPSLKGVKASVTGSS
GGIGMAVAEAPAGADVAIMYNHSPADDAEYLTGYGVKAYKCNVDPDVEKY
INLEKDFGTIDIFVANAGVAMTEGPEKFDVGRKMNKVNVDLNSVYCAHVGP
KKNKSGSLITTSMSGTIVNIPOLQAPYNAKAGVGLSKSFVEMAPFARVNSVSP
GYITLTDIADPADPEMKKMWQLTPLGREGLPQELVGAIVLYLASDAATTTGSSNIAVDG
GYTCP"
complement (3073. .3877)
/gene="VMA8"
complement (3074. .3877)
/gene="VMA8"
/feature="Vacuolar H+-ATPase subunit 8 homolog"
/codon_start=1
/transl_table=12
/product="Vma8p"
/protein_id="AAC24464.1"
/db_xref="GI:2183244"
/translation="MSGAGNREQVFPPTMTLGVKSKLKAQOQSHLLKKREALT
FRDITRIDDAKRGKGMOTAFSLAIVGATGDNISYGVOSVOKARFTVAKKON
VSGVFLTPDSHNEVDNDFKTLALAGGGOVOKAKLISKAVETLVEIASLQAPFI
LDVYIKITNRVNAIEHVITPRTNTATYINGELDENDDREPRFLTKVQEKQENANA
ABEBALAKAKAGATDELAIQDVEALDIDKADKEVDIIQEKEDVIF"

ORIGIN
Query Match 10.7%; Score 87.2; DB 8; Length 4293;
Best Local Similarity 49.7%; Pred. No. 2.9e-11;
Matches 314; Conservative 0; Mismatches 303; Indels 15; Gaps 3;

QY 182 CTGAAGATTGAAGTTGAAAGAAAGTTCCAAAGATGGCTTCAATAGCTGTGATATTCTG 241
Db 2286 CTGAATATTGAATCGAAGAAAGAAATATGCTGCAAGCCAAAGCTTAAATGATATGTAAC 2345
QY 242 ATTCGATACGGTTCACAGAGGTTTGTCTCAAGTTGCTCAAGATTTTGTGTAAGTCCAT 301
Db 2346 ATCCGAATGATGTTCTTAAATGATTAATGAAATGAAAGATTTGCTGCTATATGATA 2405
QY 302 TGCACTTGGTAAACAGAGCTGTTACTGTGAAGAACTTCCATGCAAGATTAACCAAGCA 361
Db 2406 TATTTGTTGCTAATGCTGAGATTCATGATGATGACCAAGAAATGATGTTCAAGCT 2465
QY 362 AGAAGCTGAGAGATGTAAGGTTAACTTGGGTTCTTGTATGTTTCTCAAGCT 421
Db 2466 ATGATCAATGGAAGAAAGATCGTTGATGATTAATAGAGTTTATATATGTT---GCTC 2522
QY 422 TTGTAAGCCATTTGATCAAGAAAGATTCAGAGGTCCTGTTGTTGTTGTTGTTGTTA 481
Db 2523 ATACCGTGGGACAAATCTTTAAAGATTAATCTGTTCAATTAATTAATTAATCTTCATCA 2582
QY 482 TGCTGGTGCATTTGTCACAGATCTCAAAACCAAGTTGCTACAAACATGTCGAAGGCTG 541
Db 2583 TGTCAGGACAAATGTTAATATCCCTCAATTAACAAGCTCCTTAATATGCTCTAAAGCTG 2642
QY 542 GTGTATTCATTGGCTTAAGACTTTGGCTGTGTAATGCGCTAAGTACAACATCAGATTA 601
Db 2643 CATGTACTCATTTAGCCAAATCATTTAGTGTGTAATGAGGCTAGTTGGTCTGAGATTA 2702
QY 602 ATTCCTTAACCCAGGTTACATCTACAGGTCCTTTGACCAAGATGTTATCATGTGAAC 661
Db 2703 ATTTAAATTTCTCCAGGGTATATA-----TTGACTGATATATGCTGATTTTCTGATC 2753
QY 662 AAGATTTGACACAGATGATCTCTGATATCCCAACAAGATGCGCAAGCAAG 721
Db 2754 CAGAAATGAAGAAAGAAATGCTGCAATTTGACACCTTTGGAGAGCAAGATTAACAAG 2813
QY 722 AATACATTTGCTGCTGTTTGTACTTGTCTTGAATGCTGCTTCACTAATCACTAGTG 781
Db 2814 AATTAGTGGGGCATATTATTAATCTTGGCCCTC---AAATGATCAACTTAATCACTAGTT 2870
QY 782 CCAGCTTACGTGTGAATGCTGCTTCACTTCT 813
Db 2871 CAATATTTGCTGTGATGAGGGGTTATACATGT 2902

RESULT 14
AC105434/c 126105 bp DNA linear HTG 20-MAR-2004
LOCUS Magnaporthe grisea chromosome 7 clone 18L14, *** SEQUENCING IN
DEFINITION PROGRESS ***, 2 ordered pieces.
AC105434
AC105434.2 GI:4559794
ACCESSION HTG; HTGS PHASE2.
VERSION Magnaporthe grisea (anamorph: Pyricularia grisea)
KEYWORDS Magnaporthe grisea
SOURCE Eukaryote; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
ORGANISM Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 126105)
AUTHORS Thon, M.R., Mitchell, T., Brown, D.E., Diener, S., Taro, A., Pan, H. and
Dean, R.A.
TITLE The sequence of Magnaporthe grisea chromosome 7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 126105)
AUTHORS Dean, R.A. Dr., Mitchell, T. Dr., Thon, M. Dr and Brown, D.E.
TITLE Direct Submission
JOURNAL Submitted (05-JAN-2002) Plant Pathology - Fungal Genomics
Laboratory, North Carolina State University, 840 Main Campus Drive,
Raleigh, NC 27606, USA
REFERENCE 3 (bases 1 to 126105)
AUTHORS Thon, M.R., Mitchell, T., Brown, D.E., Diener, S., Taro, A., Pan, H. and
Dean, R.A.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2004) Center for Integrated Fungal Research,

COMMENT

North Carolina State University, Plant Pathology - 840 Main Campus
Dr, Raleigh, NC 27606, USA
On Mar 20, 2004 this sequence version replaced gi:18071329.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 86195: contig of 86195 bp in length
* 86196 126105: gap of unknown length
* 86396 126105: contig of 39710 bp in length.
Location/Qualifiers
1. 126105
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/db_xref="taxon:148305"
/chromosome="7"
/clone="18L14"

FEATURES

source

ORIGIN

Query Match 10.6%; Score 86.8; DB 2; Length 126105;
Best Local Similarity 53.6%; Pred. No. 3.3e-11;
Matches 181; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 464 TTGTTTGGATTGTTCTATATGTCGTGCGATGTCAGATCTCAAAACCAAGTTGCT 523
DB 56680 TGGCTTCATCGGAGATGAGCGGCGCATGCTCAAGTCCGACGCGAGCCGCCCT 56621
QY 524 ACAACATGTCAGAGCGTGTGTATCCATTGGCTAAGACTTTGGCTTGAATGGGCTA 583
DB 56620 ACAACAGGCTCAAGCGCCCGCTGGCCACCTGGCCGCGAGCTGCGTCAAGTGGGCGC 56561
QY 584 AGTCAACATCAGATGTTAATCTTTAAACCCAGGTTACATCTACGCTCTTGAACAGA 643
DB 56580 ACGCGGCGATCGGCTCACTGATCTCTCCGCTTACATGCTACCGCCCTGACAGAA 56501
QY 644 ATGTTATCAATGTTAACAAGAAATTTTACAACAGATGATCTGTGATATCCACAACAA 703
DB 56500 AGATCTCGACGACAGCCGACCTCAAGAAAGTGAACACAGCTCATCCCGCAGGCGCA 56441
QY 704 GAATGTCGGAACCAAGAAATATCATTTGGTGTCTTTTACTTGTCTTCTGATATGCTG 763
DB 56440 GGATGGGGCGCCCAAGAGACCTGATGGGTCCCGTACCTTCTGTGTGAGCGCGTCCG 56381
QY 764 CTTCATACACTAGTGTGCGAGCTTACTGTGTGATGAT 801
DB 56380 TCTACGTACCGGCGCTGACATCGTGTGATGGCGT 56343

RESULT 15
ARS47359 927 bp DNA linear PAT 08-OCT-2004
LOCUS ARS47359
DEFINITION Sequence 2490 from patent US 6747137.
ACCESSION ARS47359
VERSION ARS47359.1 GI:53940534
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 927)
AUTHORS Weinstein, K.G. and Bush, D.
TITLE Nucleic acid sequences relating to Candida albicans for diagnostics
and therapeutics
JOURNAL Patent: US 6747137-A 2490 08-OCT-2004;
FEATURES
source Location/Qualifiers
1. 927
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 10.6%; Score 86.6; DB 6; Length 927;
Best Local Similarity 52.0%; Pred. No. 4.3e-11;
Matches 328; Conservative 0; Mismatches 279; Indels 24; Gaps 5;

QY 182 CTGAAGAAATTTGAAGTTGAAGAAAGTTCCAAAGATGGGTTCTATATGCTGTGATTTCTG 241
DB 308 CTGAATATTTTAAACCAAGACTTATAGGGTTAATCCAAAGCATTAATGTAATGTTACTG 367
QY 242 ATTCTAATCCGTTCAAGAGTGTGCTCAAGTTGCTAAGATTTTGGTAAGTCCAT 301
DB 368 ATTTCGAAGTGTGAAAGTTGTAAGTTGTCAAAACAAATTTGATCTGATTTCCGATCATGATA 427
QY 302 TGCAGTGTGTTAACACAGCTGTTACTG---TGAAACCTTCCATGTGAAGATTACCCAG 358
DB 428 TCTTTGTGCAATAGCTGGTGTGCTTCCGACCGAAGGCGCCGAATTCAGATGCAAGGAG 487
QY 359 CCAAGAACGCTGGAAGATGTGAAGTTACTTTGTGGTTTGTATGTTCTCAAG 418
DB 488 TCGACAAATGGAACAAAGTGTGATGTTAATTTAAACAGTGTATTTATTTATGCTCATG 547
QY 419 CTTTGTGTAAGCATTGATCAAGAAAGTATCAAGGGTGTCTGTGTTGTTGATTGTT 478
DB 548 TTGTTGCTCAATTTTCAAGAAAAGGTA-----AAGGTATTCATTTTCACTGCCA 601
QY 479 CTATGTCGTGCGATTTGCAACGATCTTCAAAACCAAGTTGTCTACACATGTCCAGG 538
DB 602 GTATGTGGCTTCAATTTGTAATGTCCCAATTAACAGCATTACACGCTGCTAAG 661
QY 539 CTGCTGTATTCATTGGCTTAAGACTTTGGCTGTGAAATGGGCTAAGTAAACATCAAG 598
DB 662 CTGGGGTCAAAACATTTGTCCAAATTCATAGTGTGAATGGGCAACATTTGCT--AGAG 718
QY 599 TTAATCTTTAAACCCAGGTTACATCTACGCTCTTGAACCAAGAAATGTTATCAATGTA 658
DB 719 TCAATCTGTTTCTCAAGTTACATGCTACTATCTTAATGAAATTGCTG----- 769
QY 659 ACGAAGATTTGTACAACAGATGATCTGTGATATCCACAACAAAGATGTCGAACCA 718
DB 770 ATCCGATGTCAGAGATGAATGTTGCACTTACACACTTGTGAGAGAACCAACAA 829
QY 719 AGAATATCATTTGGTGTGTTTGTACTTGTCTTCTGAAATGCTGCTTCAATACACTG 778
DB 830 GAGAGCTTGTGGTCCCTACTTGTATTTGGCTTCCGA---TGTGCAATCTTAACAACCTG 886
QY 779 GTGCAAGCTTACTGTTGATGATGTTTAC 809
DB 887 GTGCTGATCTTGTGATGATGTTGTTAC 917

Search completed: July 12, 2005, 13:24:19
Job time : 3865 secs

PT producing fermentation products or in conversion of cheap biomass to
 PT useful products.

PS Claim 6; SEQ ID NO 1; 15bp; English.

XX The present invention relates to a NADH dependent L-xylulose reductase
 CC enzyme protein and its encoding polynucleotide. The invention is useful
 CC in producing fermentation products or in conversion of cheap biomass to
 CC useful products. The present sequence is an Ambrosiozyma monospora NRRL Y
 CC -1484 NADH dependent L-xylulose reductase cDNA, ALX1 (EC 1.1.1.10).

XX Sequence 816 BP; 217 A; 159 C; 182 G; 258 T; 0 U; 0 Other;

Query Match 100.0%; Score 816; DB 12; Length 816;

Best Local Similarity 100.0%; Pred. No. 2,7e-227;

Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGACTGACTACATTCCTTCAAGTTCAGTGGCCACTTACCATTTGTCACAGTGGC 60
DB 1 ATGACTGACTACATTCCTTCAAGTTCAGTGGCCACTTACCATTTGTCACAGTGGC 60
QY 61 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 61 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 121 TTGCTTGATATCGACCAAGAAAGACCTGCTGCAAAACGCAATATACCAATACGCT 180
DB 121 TTGCTTGATATCGACCAAGAAAGACCTGCTGCAAAACGCAATATACCAATACGCT 180
QY 181 ACTGAAAGATTAAGAGTTGAAGAGTTCCAAAGATGGCTTCATATGCTGTGATATTTCT 240
DB 181 ACTGAAAGATTAAGAGTTGAAGAGTTCCAAAGATGGCTTCATATGCTGTGATATTTCT 240
QY 241 GATTCGTATACCGTTTCAAGAGTTTGTCTCAAGTTGCTAAGATTTTGTGAATGGCA 300
DB 241 GATTCGTATACCGTTTCAAGAGTTTGTCTCAAGTTGCTAAGATTTTGTGAATGGCA 300
QY 301 TTGCACTTGGTTTAAACAAGCTGTTACTGTGAAAACCTTCCCATGTGAAGATTACCAAGC 360
DB 301 TTGCACTTGGTTTAAACAAGCTGTTACTGTGAAAACCTTCCCATGTGAAGATTACCAAGC 360
QY 361 AAGAAAGCTGAGAAAGATGTGTAAGGTTAACTTTGGGTTCTTTGTATTTTCTCAAGCC 420
DB 361 AAGAAAGCTGAGAAAGATGTGTAAGGTTAACTTTGGGTTCTTTGTATTTTCTCAAGCC 420
QY 421 TTGCTTAAGCCATTGATCAAGAAGTCAAGAGGCTTCGTTGTTTGAATGGTTCT 480
DB 421 TTGCTTAAGCCATTGATCAAGAAGTCAAGAGGCTTCGTTGTTTGAATGGTTCT 480
QY 481 ATGCTGTGTCATTTGTCACAGATCTCAAAAACCAAGTTGTACAAACATGTCCAAGCT 540
DB 481 ATGCTGTGTCATTTGTCACAGATCTCAAAAACCAAGTTGTACAAACATGTCCAAGCT 540
QY 541 GGTGTATTCATTTGGCTTAAAGCTTGTGTAATGGGCTTAAGTACCAATCAGAGTT 600
DB 541 GGTGTATTCATTTGGCTTAAAGCTTGTGTAATGGGCTTAAGTACCAATCAGAGTT 600
QY 601 AATTCCTTAAACCCAGGTTACATCTAAGGCTTCTTGAACCAAGAATGTTTCAATGTTAC 660
DB 601 AATTCCTTAAACCCAGGTTACATCTAAGGCTTCTTGAACCAAGAATGTTTCAATGTTAC 660
QY 661 GAAGAATTTGACAAACAGATGATCTCTGTATCCCAACAAAGATGTCCAAACCAAG 720
DB 661 GAAGAATTTGACAAACAGATGATCTCTGTATCCCAACAAAGATGTCCAAACCAAG 720
QY 721 GAATACATTTGTGCTGTTTGTATCTGTTTGAATGCTGCTTCAATACATGCTGCT 780
DB 721 GAATACATTTGTGCTGTTTGTATCTGTTTGAATGCTGCTTCAATACATGCTGCT 780
QY 781 GCACGCTTACTGTTGATGATGTTTCACTTCTGG 816
DB 781 GCACGCTTACTGTTGATGATGTTTCACTTCTGG 816

```

RESULT 2

AAFO7518
 ID AAF07518 standard; cDNA; 3203 BP.

XX AAF07518;

DT 13-MAR-2001 (first entry)

DE Fusarium venenatum EST SEQ ID NO:41.

XX Multiple gene expression; filamentous fungal cell; EST;

KM expressed sequence tag; Fusarium venenatum; Aspergillus niger;

KM Aspergillus oryzae; Trichoderma reesei; identification; recombination;

KM culture condition; environmental stress; spore morphogenesis;

KM metabolic pathway engineering; catabolic pathway engineering; ss.

OS Fusarium venenatum.

PN WO200056762-A2.

PD 28-SEP-2000.

PF 22-MAR-2000; 2000WO-US007781.

PR 22-MAR-1999; 99US-00273623.

PA (NOVO) NOVO NORDISK BIOTECH INC.

PA (NOVO) NOVO NORDISK AS.

PI Berka RM, Rey KM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;

DR WPI; 2000-594572/56.

PT Monitoring differential expression of genes in filamentous fungal cells

PT use fluorescence-labeled nucleic acids isolated from the cells and a

PT substrate of expressed sequence tags.

PS Claim 86; Page 383-384; 3161bp; English.

CC The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FP) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FP cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FP) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring the
 CC global expression of genes from FP cells allows the production potential
 CC of the microorganisms to be improved. New genes may be discovered,
 CC possible functions of unknown open reading frames can be identified and
 CC gene copy number variation and stability can be monitored. The expression
 CC of genes can be used to study how FP cells adapt to changes in culture
 CC conditions, environmental stress, spore morphogenesis, recombination,
 CC metabolic or catabolic pathway engineering. Using ESTs provides several
 CC advantages over genomic or random cDNA clones including elimination of
 CC redundancy as one spot on an array equals one gene or open reading frame,
 CC and organisation of the microarrays based on function of the gene
 CC products to facilitate analysis of the results. AAF07478 to AAF11247
 CC represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents
 CC ESTs from Aspergillus niger; AAF11854 to AAF14678 represents ESTs from
 CC Aspergillus oryzae; and AAF14679 to AAF15337 represents ESTs from
 CC Trichoderma reesei, which are all specifically claimed in the present
 CC invention

XX Sequence 3203 BP; 770 A; 861 C; 750 G; 815 T; 0 U; 7 Other;

Query Match 12.2%; Score 99.4; DB 3; Length 3203;

Best Local Similarity 52.7%; Pred. No. 4.5e-18;

Matches 265; Conservative 0; Mismatches 231; Indels 7; Gaps 2;

```

QY 308 TTGTTAAACAGCTGTGTACTGTGAAACCTTCCATGTGAAGATTACCCAGCAAGAACG 367

```

Db 2614 TAGTACGCGGCTGGCTTCAAGAGAAATTTCAGAGCATCAACTATCCATCATGATCGTA 2673
 Qy 368 CTGAGAAAGATGTTGAAGTTAACTTTGGGTTCTTTGATGTTTCTCAAGCTTTGCTA 427
 Db 2674 TGGCGAAGTGTGGGGTGTAAATGTTGACGGTACTCTTTGGAGTTGACGTTGGCA 2733
 Qy 428 AGCCATGATCAAAAGAGTATCAAGGGCTTCTGTTGTTGATGTTGTTGTTGTTGTTG 487
 Db 2734 A--ACATCTCAATGAAAGCCAGGTAAGCTGTGATGATCTGTGTTATTTGAAACATGTCG 2790
 Qy 488 GTGCCATTGTCAACATCTCTCAAAACCAAGTTGTCTACAAACATGTCGAAGGCTGTGTTA 547
 Db 2791 GTGCTATTTGTCAATGTTCCACAGCCAGGACATATTAACGGCCAAAGGAGCTGTTTC 2850
 Qy 548 TCCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGTAAACATCAGAGTTAATTTCTT 607
 Db 2851 GTCACTGTGCTCTTCCCTCCGAGTAAAGTGGGCTCAAGCTGAAATCCGAGTCAACTGTA 2910
 Qy 608 TAAACCCAGGTTACATCAAGGCTCTTTCACCAAGAAATTTATCAATGTTACGAAGAT 667
 Db 2911 TCTCTCTGGCTACATGTTGACTGCTTACAGCAAGAAATTTCTGAACATTAACCCAGATA 2970
 Qy 668 TGTACAAAGATGATCTCTGTGATCCCAACAAGAAATGTCGAAACCAAGGAATACA 727
 Db 2971 TCGAAGAGACCTGGACATCTCTTATTTCTCAGGAGAGTATGGGACTTACTTCAAGATCTGA 3030
 Qy 728 TTGCTGCTGTTTGTACTGCTTTCTGATCTGCTGCTTCACTACATCTGCTGCTGCTGCT 787
 Db 3031 TGGGCTCTGT---GACCTTTTGTGGAGATGGCGTCTTCAATGACTGGGGGAGATC 3086
 Qy 788 TACTGCTGATGATGATGTTTCACT 810
 Db 3087 TTAGAGTATGAGAGATATACT 3109
 RESULT 3
 AAF12409
 ID AAF12409 standard; cDNA; 687 BP.
 AC AAF12409;
 XX
 DT 13-MAR-2001 (first entry)
 DE Aspergillus oryzae EST SEQ ID NO:4932.
 XX
 KM Multiple gene expression; filamentous fungal cell; EST;
 KM expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KM Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KM culture condition; environmental stress; spore morphogenesis;
 KM metabolic pathway engineering; catabolic pathway engineering; ss.
 OS Aspergillus oryzae.
 XX
 PN MO200056762-A2.
 PD 28-SEP-2000.
 XX
 PF 22-MAR-2000; 2000MO-US007781.
 XX
 PR 22-MAR-1999; 99US-00273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Beika RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;
 DR WPI; 2000-594572/56.
 PT Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags.
 XX
 PS Claim 88; Page 2078; 3161pp; English.

XX The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring the
 CC global expression of genes from FF cells allows the production potential
 CC of the microorganisms to be improved. New genes may be discovered,
 CC possible functions of unknown open reading frames can be identified and
 CC gene copy number variation and stability can be monitored. The expression
 CC of genes can be used to study how FF cells adapt to changes in culture
 CC conditions, environmental stress, spore morphogenesis, recombination,
 CC metabolic or catabolic pathway engineering. Using ESTs provides several
 CC advantages over genomic or random cDNA clones including elimination of
 CC redundancy as one spot on an array equals one gene or open reading frame,
 CC and organization of the microarrays based on function of the gene
 CC products to facilitate analysis of the results. AAF07478 to AAF11247
 CC represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents
 CC ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from
 CC Trichoderma reesei, which are all specifically claimed in the present
 CC invention
 SQ Sequence 687 BP; 152 A; 210 C; 175 G; 147 T; 0 U; 3 Other;
 Query Match 10.9%; Score 88.6; DB 3; Length 687;
 Best Local Similarity 49.8%; Pred. No. 3.3e-15;
 Matches 279; Conservative 0; Mismatches 275; Indels 6; Gaps 2;
 Qy 196 TTGAAAGAGTTCCAAAGATGGGTTCAATATCCCTGTGATTTTCTGATTTGATACCTT 255
 Db 114 TTGAAACAAATGCCCAAGTCAACGCCCACTACGCTATGTTTCCAGCTTATCTCCCTC 173
 Qy 256 CACAAGTGTGTTGCTCAAGTTGCTAAGATTTGTTGTAAGTTGCCATTGCTGGTTAA 315
 Db 174 AACGATGCCCTCTCGATATATATCTCAAGACGCAAGAT---GACACATCTGCTAC 230
 Qy 316 ACAGTGTGTTATGTTGAAATTTCCCATGTAAATTAACCAAGCCAAAGACGCTGAGAAG 375
 Db 231 TCCGCCGATTCACGAAACCTTGATGCCATCTCTACCTCAAGACCGTCTCAAAAG 290
 Qy 376 ATGGGAAGTTAACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 435
 Db 291 CTTGGGGCGTTAATGTCATGATGAACATCTTTTGCACCGGTGCGCAAGC---AC 347
 Qy 436 ATCAAGAGATATCAAGGGTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 495
 Db 348 CTCATGAGGCGCAAGTTCCGGGACGATGTCATGATGTTGATCATGTCGTGATATC 407
 Qy 496 GTCAACGATCTCTAAACCAAGTTGTTCAACATGTCGAAGGCTGTTGTTGTTGTTGTTG 555
 Db 408 GACAAAGCGCGACGCCACAGCTCTTACAAACCGCCGCAAGCGCGCTGTTGTTCACTT 467
 Qy 556 GCTAAGACTTTGGCTTGAATGGGCTAAGTAAACATCAAGTTAATTTCTTAAACCA 615
 Db 468 GCCGCTACTTCCCGCGAATGGCGGCTACACATCTCGGGTACGTGATAGACCTT 527
 Qy 616 GATTACATCTAAGCTCTTTGACCAAGATGTTATCAATGTTACGAAGATTGTACAA 675
 Db 528 CGATACATGCTTACTGCGCTGACCGCAAGTTTGTGATGAACCCCGCATTCGGGAC 587
 Qy 676 AGATGATCTCTGATATCCCAACAAGATGTCGAACCAAGAAATATCATTTGTTGCT 735
 Db 588 AAGTGATCTGCTCATCTCCACCGGCAAGTGGTACTCCCGAGGACCTGATGGTCTCC 647
 Qy 736 GTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 755
 Db 648 GTTACCTTTCTGCTCAAGGA 667

RESULT 4
 ADS61049
 ID ADS61049 standard; cDNA; 637 BP.
 AC ADS61049;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #13036.
 XX
 KM Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US200323675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 36723; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transforming plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 637 BP; 149 A; 177 C; 169 G; 142 T; 0 U; 0 Other;

Query Match 10.0%; Score 81.8; DB 13; Length 637;
 Best Local Similarity 55.0%; Pred. No. 3.1e-13;

Matches 183; Conservative 0; Mismatches 147; Indels 3; Gaps 1;
 QY 304 CACTTGATTAAACAGCTGGTTACTGTGAAAACTTCCATGTGAAAGATTACCGCAAG 363
 DB 70 CATCTGTCACGTCGTGGGATTTACGAAACTTTCAGCCCATACCATACGAC 129
 QY 364 AACGCTGAGAAATGTTGAAGTTAACTTTGGGTTCTTGTATGTTCTCAAGCTTT 423
 DB 130 CGTATGACAGAAAGCTTGGGGGGTGAACGTGACGACATATCTGTTGCTACGGCGTG 189
 QY 424 GTTAAGCATTTATCAAGAAAGGTATCAAGGGTCTTCTGTTGTTTGTGATTGTTCTATG 483
 DB 190 GCAAGGACTTGAAT--GAGGCCAAGAGCCGCCGAGATTTGTGATGATCGGTAGATG 246
 QY 484 TCTGTGTCATTTGCTCAACGATTCCTCAAAACAAAGTTGTCTCAACATGTCGAAGCTGT 543
 DB 247 TCTGTGTCATTTGCTCAACGATTCCTCAAAACAAAGCTGTCAACAGCCGCAAGGCGCC 306
 QY 544 GTTATCCATTTGGCTTAAGACTTTGGCTTGTGATTTGGGCTTAAGTCAACATCAGATTAT 603
 DB 307 GTTCGCCACCTTCTGCTTCTTTCGCGGTGAATGGGCAAGATGATGATCAGGGTAAAC 366
 QY 604 TCTTTAAACCCAGGTTATCATCTACGGTCTTTG 636
 DB 367 TGCATTAGCCCTGATATCATCTGACTGCTCTG 399

RESULT 5
 ADS48561
 ID ADS48561 standard; cDNA; 735 BP.
 XX
 AC ADS48561;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #3304.
 XX
 KM Recombinant DNA construct; transformed plant; improved plant property;
 KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KM bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US200323675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 26991; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to

XX Multiple gene expression; filamentous fungal cell; EST;
 KW expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;
 KW *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 XX
 OS *Aspergillus oryzae*.
 XX
 PN WO200056762-A2.
 XX
 PD 28-SEP-2000.
 XX
 PF 22-MAR-2000; 2000WO-US007781.
 XX
 PR 22-MAR-1999; 99US-00273623.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
 DR WPI; 2000-594572/56.
 XX
 PT Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags.
 XX
 PS Claim 88, Page 2039, 3161pp; English.
 XX
 OS The present invention describes a method for monitoring differential
 CC expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring the
 CC global expression of genes from FF cells allows the production potential
 CC of the microorganisms to be improved. New genes may be discovered,
 CC possible functions of unknown open reading frames can be identified and
 CC gene copy number variation and stability can be monitored. The expression
 CC of genes can be used to study how FF cells adapt to changes in culture
 CC conditions, environmental stress, spore morphogenesis, recombination,
 CC metabolic or catabolic pathway engineering. Using ESTs provides several
 CC advantages over genomic or random cDNA clones including elimination of
 CC redundancy as one spot on an array equals one gene or open reading frame,
 CC and organization of the microarrays based on function of the gene
 CC products to facilitate analysis of the results. AAF07478 to AAF11247
 CC represents ESTs from *Fusarium venenatum*; AAF11248 to AAF11853 represents
 CC ESTs from *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from
 CC *Aspergillus oryzae*; and AAF14879 to AAF15337 represents ESTs from
 CC *Trichoderma reesei*, which are all specifically claimed in the present
 CC invention
 CC
 SQ Sequence 1146 BP; 262 A; 313 C; 274 G; 297 T; 0 U; 0 Other;
 XX
 XX
 Query Match 9.1%; Score 74.2; DB 3; Length 1146;
 Best Local Similarity 50.6%; Pred. No. 6, 6e-11;
 Matches 206; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 529 ATGTCCAGGCTGTGTATTCATTGCTTAAGACTTTGGCTTGTGATGGCTTAAGTAC 588
 |||||
 Db 744 GTTAGACAGCTTCAGTCTTCACTTCCCGTAACTCCCATGTGAGTGGGCCGTAC 803
 QY 589 AACATCAGATTAATCTTAAACCAAGTTACATCAGGTCCTTGTGACCAAGATGT 648
 |||||
 Db 804 AACATTCAGTCAACACCATCTGCGCGCTCAATTGTACTGCCATGTGTGAGAACTC 863
 QY 649 ATCAATGTGTAACGAAATTTGTACAACAGATGATCTTGTATCCCAACAAAGATG 708
 |||||
 Db 864 TTGCTTAGTTCCTTGAAGCGTCGCGAAGATGCGCCAAACATTAACATGCTGGAGCTGT 923
 QY 709 TCGAACCAAGAAATACATTGCTGCTTTTGTGACTTCTTCTGA 755
 |||||
 Db 924 TCTACCCCTTAACGAGTACCGTGGCGTCCGCTCTTCTTCAGTGA 970

RESULT 8
 ADS47380
 ID ADS47380 standard; cDNA; 783 BP.
 AC ADS47380;
 XX
 XX 02-DEC-2004 (first entry)
 DT
 XX
 DE Bacterial polynucleotide #2123.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.
 XX
 PN US200323675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINKLE/) HINKLE G J.
 PA (SLATER/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 25810; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production, this sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 783 BP; 193 A; 186 C; 173 G; 231 T; 0 U; 0 Other;

Query Match 8.3%; Score 67.6; DB 13; Length 783;
Best Local Similarity 51.2%; Pred. No. 4.7e-09;
Matches 213; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

QY 171 CAATAACGCTACTGAGAAATTGAAAGTTGCAAGAGTGGTTCATATGCTTG 230
DB 171 CAAAAGGCTGCAAGAAATTGCCAGGCTAATGCTGCAAGACTTATACATGCAAGTG 230
QY 221 TGAATTTCTGATTCGATACCGTTCAAGAGTGGTTCAGAGTTGCTAAGATTGG 290
DB 221 CGATGTCATATTCCTAAAGAGGTAGAACATGCTTGTGAGATTCAAAAGGCTTTGA 290
QY 231 TAAAGTTCAGCTTGAAGCTTGTAAACACAGCTGTTACTGTGAAAACCTCCATGTGAGA 350
DB 231 TACTATATGATATCGTTGTGCCAACAACGTAATTTGCACTGGCAAGTCGGCATTT---GA 347
QY 351 TTACCAAGCAAGAAAGCTGAGAAAGTGTGAAGGTTAACTTGGTGTCTTTGTATGT 410
DB 348 CATGACTTACGAAGAAATTGCTAATGAATTAACGTAACCTGCTGGGTCTTCAACGT 407
QY 411 TTCTAGAGCTTGTGTAAGCCATGATCAAGAGTATCAAGGTCCTTCTGTTT 470
DB 408 CGCTACAAATGCTGGGCCCATCTTCCAAAAGCAAG-----CAATGGAGCCTTGTATAC 461
QY 471 GATTTGTTCTATGTCTGTGTCATATGTCATCAAGATCCTCAAAACCAAGTTGTCTACACAT 530
DB 472 TACTGCTTCATATGCTGTGTGTGTTGTTTACGTTCTCAACAACATGTCCTTACACAC 521
QY 531 GTCCAGAGGTGTGTATTCATTGCTGAAGACTTGTGGTGTGTGAATGGGCTAAGT 586
DB 522 TTCCAAAGGCGGTGTCATTCATCACTCAAGAGTTTGGAGTGCATGGCGCAAT 577

RESULT 9
ADN62512
ID ADN62512 standard; DNA; 789 BP.

XX ADN62512;

XX 01-JUL-2004 (first entry)

XX A. thaliana gene SEQ ID NO:17.

XX ds; gene; seed size; organ size; plant; transgenic.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

XX FT CDS 1..789 a

XX WO2003096797-A2.

XX 27-NOV-2003.

XX 14-MAY-2003; 2003MO-US014989.

XX 15-MAY-2002; 2002US-0381100P.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX PI Fernandes M, Xie Z, Dotson SB;

XX DR WPI: 2004-02481/04.

XX DR P-PSDB; ADN62513.

XX PT Increasing seed and organ size of a plant by transforming the plant with
XX a DNA construct comprising a promoter that functions in plants and
XX PT selecting a desired plant from a population of transformed plants
XX containing the DNA construct.

XX PS Claim 1; SEQ ID NO 17; 71pp; English.

XX SQ The invention relates to a novel method for increasing seed and organ
XX size of a plant by: transforming the plant with a DNA construct
XX comprising a promoter that functions in plants, operably linked to a DNA
XX molecule that encodes a protein; and selecting a desired plant from a
XX CC population of transformed plants containing the DNA construct; where the
XX CC desired plant exhibits increased seed and organ size compared to a plant
XX CC of a same plant species not transformed to contain the DNA construct. The
XX method of the invention is useful for increasing seed and organ size of a
XX plant. The present sequence is used in the exemplification of the
XX invention.

XX SQ Sequence 789 BP; 220 A; 151 C; 183 G; 235 T; 0 U; 0 Other;

Query Match 8.1%; Score 66.4; DB 12; Length 789;
Best Local Similarity 56.5%; Pred. No. 1.1e-08;
Matches 165; Conservative 0; Mismatches 121; Indels 6; Gaps 2;

QY 519 TGTCTACAACATGTCGAAGCTGTGTGTTATCCATTTGGCTAAGACTTTGGCTTGAATG 578
DB 471 TATCTATATGTTAAGAAAGAGCTTGTATCACTACACAAAACCTTGGAGTGAAG 530
QY 579 GGCTAAGTACACATCAGATTAATCTTTAAACCAAGTTACATCAAGCTCTTGTAC 638
DB 531 GGCAAGATATGACATTAAGAGCCAACTGTTGCTCTATTTTATCTACATGCTATGGC 590
QY 639 CAGAAATGTTATCATATGTTAAGAAATTGTACAAAGATGATCTGTGATCCCA 698
DB 591 TCAACCTTTTCAAG---ACGCCGATTACAGAAAGTTGGTGAAGAACTTCACT 647
QY 699 ACAAGAAATGTCGAACCAAGAAATATACATTTGCTGTTTGTACTTGTGAATC 758
DB 648 TGGTGGCGCTGGAGAGCCAAATGAG---TTTCATCACTTGTGGCTTTTGTGTCTAC 704
QY 759 TGTGCTTCATACACTACTGTGTGTCAGCTTACTGTGTATGATGGTGTCTACT 810
DB 705 TGCAGCTTCATATATTTACTGTGTCAGACCATTTGTTGATGGAGTCTCACT 756

RESULT 10
ADS48180
ID ADS48180 standard; cDNA; 774 BP.

XX ADS48180;

XX 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #2923.

XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.

XX Bacteria.

| | | |
|----|---|--|
| XX | US2003233675-A1. | |
| XX | 18-DEC-2003. | |
| XX | 20-FEB-2003; 2003US-00369493. | |
| PF | 21-FEB-2002; 2002US-0360039P. | |
| XX | | |
| PA | (CAOY/) CAO Y. | |
| PA | (HINK/) HINKLE G J. | |
| PA | (SLAT/) SLATER S C. | |
| PA | (CHEN/) CHEN X. | |
| PA | (GOLD/) GOLDMAN B S. | |
| XX | | |
| PI | Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS; | |
| DR | WPI; 2004-061375/06. | |
| XX | | |
| PT | New recombinant DNA construct comprising a promoter positioned to provide | |
| PT | for expression of a polynucleotide encoding a polypeptide from a | |
| PT | microbial source, useful for producing plants with improved properties. | |
| XX | | |
| PS | Claim 1; SEQ ID NO 26610; 122pp; English. | |
| XX | | |
| CC | The invention relates to a recombinant DNA construct comprising a | |
| CC | promoter functional in a plant cell, where the promoter is positioned to | |
| CC | provide for expression of a polynucleotide encoding a polypeptide from a | |
| CC | microbial source. The invention also relates to a transformed plant | |
| CC | comprising the recombinant DNA construct and a method of producing a | |
| CC | transformed plant having an improved property. The plant is a crop plant | |
| CC | such as maize or soybean. The method of producing a transformed plant | |
| CC | having an improved property comprises transforming a plant with the | |
| CC | recombinant DNA construct and growing the transformed plant, where the | |
| CC | polynucleotide or polypeptide is useful for improving plant properties. | |
| CC | The recombinant DNA construct is useful for producing plants with | |
| CC | improved plant properties, e.g. improved cold, heat or drought tolerance, | |
| CC | tolerance to herbicides, extreme osmotic conditions, pathogens or pests, | |
| CC | increased resistance to plant disease, better growth rate by modification | |
| CC | of the cell cycle pathway with plant growth regulators, increased rate of | |
| CC | homologous recombination, modified seed oil or protein yield and/or | |
| CC | content, improved yield by modification of carbohydrate, nitrogen or | |
| CC | phosphorus use and/or uptake, by modification of photosynthesis or by | |
| CC | providing improved plant growth and development under at least one stress | |
| CC | condition, improved lignin production or improved galactomanan | |
| CC | production. This sequence represents a bacterial polynucleotide used in | |
| CC | the scope of the invention. Note: The sequence data for this patent did | |
| CC | not form part of the printed specification but was obtained in electronic | |
| CC | format from USPTO at seqdata.uspto.gov/sequence.html . | |
| XX | | |
| XX | Sequence 774 BP; 226 A; 160 C; 229 G; 159 T; 0 U; 0 Other; | |
| XX | | |
| XX | Query Match 7.8%; Score 63.6; DB 13; Length 774; | |
| XX | Best Local Similarity 61.4%; Pred. No. 6.9e-08; | |
| XX | Matches 102; Conservative 0; Mismatches 64; Indels 0; Gaps 0 | |
| QY | 472 ATGCGTTTATGTCGTGTCGCAATGTCGACGATCTCTCAAAACCAAGTTGTTCAACAATG 531 | |
| DB | 436 ATCGCGTCATGTCGCGACACATCGGAACAAACCTCAGAAAGCAGACAGTTAAACAGCT 495 | |
| QY | 532 TCCAAAGCTGCGTTTATTCATTTGGCTAAGACCTTTGGCTTTGGAAGGGGTAAAGTACAC 591 | |
| DB | 496 TCGAAGCGCGGTGTGATCATCTCAACAGATCTCTGCGCGCGCAGTGGGCGCCGTACGGA 555 | |
| QY | 592 ATCAGAGTTAATCTTTAAACCAAGGTTACATCTAAGGTCCTTTGA 637 | |
| DB | 556 ATCAGGTTGAACAGCATTAAGCCCCGATACATAGAAACACCTTCA 601 | |
| XX | | |
| XX | RESULT 11 | |
| XX | AAFL1343 | |
| XX | ID AAFL1343 standard; cDNA; 636 BP. | |
| AC | AAFL1343; | |

XX 13-MAR-2001 (first entry)
 DT
 XX Aspergillus niger EST SEQ ID NO:1866.
 DE
 XX Multiple gene expression; filamentous fungal cell; EST;
 XX expressed sequence tag; Fusarium venenatum; Aspergillus niger;
 KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
 KW culture condition; environmental stress; spore morphogenesis;
 KW metabolic pathway engineering; catabolic pathway engineering; ss.
 XX
 XX Aspergillus niger.
 OS
 XX
 XX WO200056762-A2.
 PN
 XX
 XX 28-SEP-2000.
 PD
 XX
 XX 22-MAR-2000; 2000WO-US007781.
 PF
 XX
 XX 22-MAR-1999; 99US-00273623.
 PR
 XX
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO NORDISK AS.
 XX
 XX Berka RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;
 PI WPI; 2000-594572/56.
 DR
 XX
 XX Monitoring differential expression of genes in filamentous fungal cells
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a
 PT substrate of expressed sequence tags.
 XX
 PS Claim 87; Page 1728; 3161pp; English.
 XX
 XX The present invention describes a method for monitoring differential
 XX expression of genes in a first filamentous fungal (FF) cell relative to
 CC expression of the same genes in one or more second filamentous fungal
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
 CC are used in the methods for monitoring differential expression of genes
 CC in a first filamentous fungal (FF) cell relative to expression of the
 CC same genes in one or more second filamentous fungal cells. Monitoring the
 CC global expression of genes from FF cells allows the production potential
 CC of the microorganisms to be improved. New genes may be discovered,
 CC and possible functions of unknown open reading frames can be identified,
 CC and gene copy number variation and stability can be monitored. The expression
 CC of genes can be used to study how FF cells adapt to changes in culture
 CC conditions, environmental stress, spore morphogenesis, recombination,
 CC metabolic or catabolic pathway engineering. Using ESTs provides several
 CC advantages over genomic or random cDNA clones including elimination of
 CC redundancy as one spot on an array equals one gene or open reading frame,
 CC and organisation of the microarrays based on function of the gene
 CC products to facilitate analysis of the results. AA07478 to AAF11247
 CC represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents
 CC ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from
 CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from
 CC Trichoderma reesei, which are all specifically claimed in the present
 CC invention
 CC
 XX
 XX Sequence 636 BP; 169 A; 155 C; 161 G; 149 T; 0 U; 2 Other;
 SQ
 XX
 XX Query Match 7.4%; Score 60.6; DB 3; Length 636;
 XX Best Local Similarity 51.3%; Pred. No. 4.7e-07;
 XX Matches 194; Conservative 0; Mismatches 175; Indels 9; Gaps 2
 XX
 XX 211 AAGATGGGTTCAATGATCCGTGATATTTCTGATTCGATACCGTTACAGAGTGTTGCT 270
 DB 200 AAGTCAGCTGCTATTAAGACTATATGTCCGAATTCAGAGAGATTCAGAGGCTAATCCAG 259
 XX
 XX 271 CAAGTTGCTAAGATTTTGGTAAGTGGCATTGCCATTGGTTAACAACAGCTGTTACTGT 330
 DB 260 CAGATTGCTTCAAGCTTTGGCAGAGCTGTGATATTATTCGTGTGAAC---TCGGGGGCTCAC 316

PS Claim 1; Page 10-11; 14pp; Japanese.
XX
CC The invention relates to gene encoding D-arabinitol dehydrogenase.
CC isolated from *Bacillus* sp. IKD-5A868 strain. The protein can be expressed
CC by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC as a clinical diagnosing agent for mycosis. The present sequence
CC represents the coding sequence of the D-arabinitol dehydrogenase
SQ Sequence 774 BP; 155 A; 214 C; 271 G; 134 T; 0 U; 0 Other;
Query Match 7.2%; Score 59; DB 3; Length 774;
Best Local Similarity 50.4%; Pred. No. 1.5e-06;
Matches 171; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 478 TCTATGTCGTGGCCATTGTCACAGATCTTCAAAACCAAGTTGCTTACATGTCACAG 537
DB 439 TCCATGTCGGGCGCTGATGTCATATACCGCGACGCGGCGGCTTCAATGTCCTCAAG 498
QY 538 GCTGTGTATTCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGTACATCAGA 597
DB 499 GCGGGGCTCATTTATGCTGACCAAGAGCTGGCATCCAAATGGGCGCGCATGGGCTGGC 558
QY 598 GTTATTTCTTTAAACCCAGTTACATCTACGCTCTTGTGACCAAGATGTTATCAATGT 657
DB 559 GTCAACACGATTTGGCGCCGGCTACATGAAGACGAGCGAACCGTATTTGCGCGCC 618
QY 658 AACGAGAAATTGTACACAGATGATCTCTGTGATCCCAACAAAGATGTCGAACCA 717
DB 619 GAGGCGAGATGATGACAGAGTGCTGCTATGACCCGATGAGGCGCGCGCTTCC- 677
QY 718 AAGGAATACATTGGGCTGTTTGTACTGCTTCTGATCTGATCTGCTTCAATACACTACT 777
DB 678 --GACGAGCTGGGCGGCGATTCGCTTACCTTCTTGAGACGCTTCTCTTGGCCACG 735
QY 778 GATGCCAGCTTACTGTTGATGATGTTGATTTCACTTCTTGG 816
DB 736 GCGGCGGTGTTACGATTTGACGCGGGGCTATACGATTTGG 774
RESULT 14
AAZ46763
ID AAZ46763 standard; DNA; 2255 BP.
XX
AC AAZ46763;
XX
DT 31-MAR-2000 (first entry)
XX
DE *Bacillus* D-arabinitol dehydrogenase encoding genomic DNA.
XX
KM D-arabinitol dehydrogenase; clinical diagnosis; mycosis; ds.
XX
OS *Bacillus* sp.
XX
XX
XX Key Location/Qualifiers
FH CDS 366..1142
FT /tag= a
FT /product= "D-arabinitol dehydrogenase"
FT /note= "the coding sequence is also given in AAZ46762"
XX
PN JP1332569-A.
XX
XX 07-DEC-1999.
PD
XX
PF 26-MAY-1998; 98JP-00143637.
XX
PR 26-MAY-1998; 98JP-00143637.
XX
XX (IKED-) IKEDA SHOKKEN KK.
PA (NIPK) NIPON KAYAKU KK.
XX
XX MPI: 2000-091353/08.
DR P-PSDB; AA156815.
XX

PT Arabinitol dehydrogenase gene encoding D-arabinitol dehydrogenase -
PT useful as a clinical diagnosing agent for mycosis.
XX
PS Claim 6; Page 12; 14pp; Japanese.
XX
CC The invention relates to gene encoding D-arabinitol dehydrogenase.
CC isolated from *Bacillus* sp. IKD-5A868 strain. The protein can be expressed
CC by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC as a clinical diagnosing agent for mycosis. The present sequence
CC represents the DNA encoding the D-arabinitol dehydrogenase
SQ Sequence 2255 BP; 418 A; 623 C; 728 G; 486 T; 0 U; 0 Other;
Query Match 7.2%; Score 59; DB 3; Length 2255;
Best Local Similarity 50.4%; Pred. No. 2.4e-06;
Matches 171; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 478 TCTATGTCGTGGCCATTGTCACAGATCTTCAAAACCAAGTTGCTTACATGTCACAG 537
DB 804 TCCATGTCGGGCGCTGATGTCATATACCGCGACGCGGCGGCTTCAATGTCCTCAAG 863
QY 538 GCTGTGTATTCATTGGCTAAGACTTTGGCTTGTGAATGGGCTAAGTACATCAGA 597
DB 864 GCGGGGCTCATTTATGCTGACCAAGAGCTGGCATCCAAATGGGCGCGCATGGGCTGGC 923
QY 598 GTTATTTCTTTAAACCCAGTTACATCTACGCTCTTGTGACCAAGATGTTATCAATGT 657
DB 924 GTCAACACGATTTGGCGCCGGCTACATGAAGACGAGCGAACCGTATTTGCGCGCC 983
QY 658 AACGAGAAATTGTACACAGATGATCTCTGTGATCCCAACAAAGATGTCGAACCA 717
DB 984 GAGGCGAGATGATGACAGAGTGCTGCTATGACCCGATGAGGCGCGCGCTTCC- 1042
QY 718 AAGGAATACATTGGGCTGTTTGTACTGCTTCTGATCTGATCTGCTTCAATACACTACT 777
DB 1043 --GACGAGCTGGGCGGCGATTCGCTTACCTTCTTGAGACGCTTCTCTTGGCCACG 1100
QY 778 GATGCCAGCTTACTGTTGATGATGTTGATTTCACTTCTTGG 816
DB 1101 GCGGCGGTGTTACGATTTGACGCGGGGCTATACGATTTGG 1139
RESULT 15
AAC42189
ID AAC42189 standard; DNA; 798 BP.
XX
AC AAC42189;
XX
DT 17-OCT-2000 (first entry)
XX
DE *Arbidiopsis* thaliana DNA fragment SEQ ID NO: 34619.
XX
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX
OS *Arbidiopsis* thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
PP 25-FEB-2000; 2000EP-00301439.
XX
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
XX

PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 30-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 04-MAY-1999; 99US-0132407P.
 PR 05-MAY-1999; 99US-0132484P.
 PR 06-MAY-1999; 99US-0132485P.
 PR 07-MAY-1999; 99US-0132486P.
 PR 11-MAY-1999; 99US-0132863P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139111P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143547P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 21-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 23-JUL-1999; 99US-0145129P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145221P.
 PR 26-JUL-1999; 99US-0145224P.
 PR 27-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145919P.
 PR 02-AUG-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 05-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149175P.
 PR 17-AUG-1999; 99US-0149168P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159283P.
 PR 13-OCT-1999; 99US-0159284P.
 PR 13-OCT-1999; 99US-0159285P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 12:10:28 ; Search time 185 Seconds

(without alignments)
7217.307 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816
Sequence: 1 atgactgactcattccac.....atgctgcttcattcttgg 816

Scoring table:

IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1202764 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfileseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|---------------------|--------------------|
| 1 | 238.2 | 29.2 | 696 | US-08-336-198C-6 | Sequence 6, Appli |
| 2 | 238.2 | 29.2 | 696 | US-09-184-965-6 | Sequence 6, Appli |
| 3 | 218.6 | 26.8 | 846 | US-09-248-796A-29 | Sequence 29, Appli |
| 4 | 86.6 | 10.6 | 927 | US-09-248-796A-2490 | Sequence 2490, Ap |
| 5 | 74.6 | 9.1 | 852 | US-09-734-237B-71 | Sequence 71, Appli |
| 6 | 61.8 | 7.6 | 873 | US-09-248-796A-2654 | Sequence 2654, Ap |
| 7 | 55 | 6.7 | 336 | US-09-248-796A-2489 | Sequence 2489, Ap |
| 8 | 53 | 6.5 | 372 | US-09-248-796A-6987 | Sequence 6987, Ap |
| 9 | 53 | 6.5 | 774 | US-09-134-001C-725 | Sequence 725, App |
| 10 | 51.8 | 6.3 | 825 | US-09-248-796A-730 | Sequence 2730, Ap |
| 11 | 51.2 | 6.3 | 852 | US-09-367-012-2 | Sequence 2, Appli |
| 12 | 51.2 | 6.3 | 852 | US-09-777-157A-2 | Sequence 2, Appli |
| 13 | 51.2 | 6.3 | 852 | US-09-734-237B-69 | Sequence 69, Appli |
| 14 | 51.2 | 6.3 | 951 | US-09-489-039A-5819 | Sequence 5819, Ap |
| 15 | 51.2 | 6.3 | 1303 | US-09-367-012-8 | Sequence 8, Appli |
| 16 | 51.2 | 6.3 | 1303 | US-09-777-157A-8 | Sequence 8, Appli |
| 17 | 50.4 | 6.2 | 4411529 | US-09-103-840A-1 | Sequence 1, Appli |
| 18 | 49.8 | 6.1 | 873 | US-09-248-796A-2489 | Sequence 2489, Ap |
| 19 | 48.2 | 5.9 | 3993 | US-09-710-279-3985 | Sequence 3985, Ap |
| 20 | 47.4 | 5.8 | 4403765 | US-09-103-840A-2 | Sequence 2, Appli |
| 21 | 47 | 5.8 | 792 | US-09-134-001C-1675 | Sequence 1675, Ap |
| 22 | 45.2 | 5.5 | 2774 | US-09-363-189B-5 | Sequence 5, Appli |
| 23 | 44.4 | 5.4 | 1248 | US-10-024-806-7 | Sequence 7, Appli |
| 24 | 43 | 5.3 | 879 | US-09-468-738A-1 | Sequence 1, Appli |
| 25 | 43 | 5.3 | 879 | US-09-940-019-1 | Sequence 1, Appli |
| 26 | 43 | 5.3 | 879 | US-09-940-037A-1 | Sequence 1, Appli |
| 27 | 43 | 5.3 | 891 | US-09-468-738A-22 | Sequence 22, Appli |

| | | | | | |
|----|------|-----|-------|---------------------|--------------------|
| 28 | 43 | 5.3 | 891 | US-09-940-019-22 | Sequence 22, Appli |
| 29 | 43 | 5.3 | 891 | US-09-940-037A-22 | Sequence 22, Appli |
| 30 | 43 | 5.3 | 1787 | US-09-468-738A-17 | Sequence 17, Appli |
| 31 | 43 | 5.3 | 1787 | US-09-940-019-17 | Sequence 17, Appli |
| 32 | 43 | 5.3 | 1787 | US-09-940-037A-17 | Sequence 17, Appli |
| 33 | 42.2 | 5.2 | 789 | US-09-710-279-3157 | Sequence 3157, Ap |
| 34 | 42.2 | 5.2 | 3210 | US-09-710-279-3922 | Sequence 3922, Ap |
| 35 | 42.2 | 5.2 | 3368 | US-09-710-279-3922 | Sequence 4169, Ap |
| 36 | 41.4 | 5.1 | 825 | US-09-134-001C-1594 | Sequence 1594, Ap |
| 37 | 41.4 | 5.1 | 916 | US-09-710-279-381 | Sequence 381, App |
| 38 | 41.4 | 5.1 | 3024 | US-09-710-279-3426 | Sequence 3426, Ap |
| 39 | 41.2 | 5.0 | 2472 | US-09-248-796A-148 | Sequence 148, App |
| 40 | 41 | 5.0 | 437 | US-09-468-738A-14 | Sequence 14, Appli |
| 41 | 41 | 5.0 | 437 | US-09-940-019-14 | Sequence 14, Appli |
| 42 | 41 | 5.0 | 437 | US-09-940-037A-14 | Sequence 14, Appli |
| 43 | 41 | 5.0 | 837 | US-09-328-352-2671 | Sequence 2671, Ap |
| 44 | 40.8 | 5.0 | 601 | US-09-949-016-62984 | Sequence 62984, A |
| 45 | 40.8 | 5.0 | 50263 | US-09-949-016-13563 | Sequence 13563, A |

ALIGNMENTS

RESULT 1
US-08-336-198C-6
; Sequence 6, Application US/08336198C
; Patent No. 5866382
; GENERAL INFORMATION:
; APPLICANT: Hallborn, Johan
; APPLICANT: Penttila, Merja
; APPLICANT: Ojamo, Heikki
; APPLICANT: Keranen, Sirikka
; APPLICANT: Hahn-Hagedorn, Barbel
; APPLICANT: Waldfriedson, Mats
; APPLICANT: Aikarsinen, Ulla
; TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336, 198C
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 base pairs
; TYPE: nucleic acid
; STRANDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Pichia stipitis
; STRAIN: CBS-6054
; FEATURE:
; NAME/KEY: CDS

LOCATION: 1..693
OTHER INFORMATION: /standard name="xyli1col
OTHER INFORMATION: dehydrogenase"
PUBLICATION INFORMATION:
DOCUMENT NUMBER: FI 901771
FILING DATE: 06-APR-1990
US-08-336-198C-6

Query Match 29.2%; Score 238.2; DB 2; Length 696;
Best Local Similarity 64.4%; Pred. No. 9.4e-65;
Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 227 CCTGTGATATTTCTGATTTCTGATACCGTTCAACAGCTGTTGCTCAAGTGTCTAAGTT 286
DB 107 CCTGCAACATTTGGGGATGCTGAGCACTAATTTCAAGTGTCTAAGTTCAACAGTACC 166
QY 287 TTGTGATGTTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 346
DB 167 ACGGCAAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 226
QY 347 AAGATTACCCAGCCAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406
DB 227 AAGCTTACCCAGCCAG 286
QY 407 ATGTTTCTCAAGCTTTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 466
DB 287 AAGTTTCTCAAGCTTTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 343
QY 467 TTTTATGTTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 526
DB 344 TTTTATGTTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 403
QY 527 ACATGTCCAGAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 586
DB 404 ACATGTCCAGAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 463
QY 587 ACAATCATGAGTTAATTTTAAACCCAGGTTAATCTAAGTCTTGTGACCAAGATG 646
DB 464 ACAATCATGAGTTAATTTTAAACCCAGGTTAATCTAAGTCTTGTGACCAAGATG 523
QY 647 TTTATCAAGTGTACCAAGATTTGTACCAAGATTTGTACCAAGATTTGTACCAAG 706
DB 524 TTTATCAAGTGTACCAAGATTTGTACCAAGATTTGTACCAAGATTTGTACCAAG 583
QY 707 TTTTCCAGCAAGAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 766
DB 584 TTTTCCAGCAAGAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 643
QY 767 CATACATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 805
DB 644 CATACATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682

RESULT 2
US-09-184-965-6
Sequence 6, Application US/09184965
Patent No. 6582944
GENERAL INFORMATION:
APPLICANT: Hallborn, Johan
APPLICANT: Penttila, Merja
APPLICANT: Ojamo, Heikki
APPLICANT: Keranen, Sirkka
APPLICANT: Hahn-Hagerdal, Barbel
APPLICANT: Waldfriedson, Mats
APPLICANT: Airaksinen, Ulla
TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Birch, Stewart, Kolaesch & Birch
STREET: 301 N. Washington St.
CITY: Falls Church
STATE: Virginia
COUNTRY: USA

ZIP: 22046-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/184,965
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,198
FILING DATE: 03-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 696 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Pichia stipitis
STRAIN: CBS-6054
FEATURE:
NAME/KEY: CDS
LOCATION: 1..693
OTHER INFORMATION: /standard name="xyli1col
OTHER INFORMATION: dehydrogenase"
PUBLICATION INFORMATION:
DOCUMENT NUMBER: FI 901771
FILING DATE: 06-APR-1990
US-09-184-965-6

Query Match 29.2%; Score 238.2; DB 4; Length 696;
Best Local Similarity 64.4%; Pred. No. 9.4e-65;
Matches 373; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 227 CCTGTGATATTTCTGATTTCTGATACCGTTCAACAGTGTGTTGCTCAAGTGTCTAAGATT 286
DB 107 CCTGCAACATTTGGGGATGCTGAGCACTAATTTCAAGTGTGTTGCTCAAGTGTCTAAG 166
QY 287 TTGTGATGTTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 346
DB 167 ACGGCAAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 226
QY 347 AAGATTACCCAGCCAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 406
DB 227 AAGCTTACCCAGCCAG 286
QY 407 ATGTTTCTCAAGCTTTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 466
DB 287 AAGTTTCTCAAGCTTTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 343
QY 467 TTTTATGTTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 526
DB 344 TTTTATGTTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 403
QY 527 ACATGTCCAGAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 586
DB 404 ACATGTCCAGAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 463
QY 587 ACAATCATGAGTTAATTTTAAACCCAGGTTAATCTAAGTCTTGTGACCAAGATG 646
DB 464 ACAATCATGAGTTAATTTTAAACCCAGGTTAATCTAAGTCTTGTGACCAAGATG 523

QY 647 TTATCATGTGTAAGAAATTTGATACAGATGATCTGTGTATCCCAACAAGAA 706
DB 524 TATATTTCTGGCCACACAGATGAAGAAAGCTGGGAATCCAGATCCCATGAAGACA 583
QY 707 TGTCCGAACCAAGAAATATCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 766
DB 584 TGGCCGAACCAAGAAATCTGTGGGTCCATCTTATCTTGGCAAGCGAGATGCTTCTT 643
QY 767 CATACACTACTGT 805
DB 644 CCTACACTACTGT 682

RESULT 3
US-09-248-796A-29

/ Sequence 29, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ LENGTH: 846
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-29

Query Match 26.8%; Score 218.6; DB 4; Length 846;
Best Local Similarity 62.0%; Pred. No. 1.6e-58;
Matches 385; Conservative 0; Mismatches 224; Indels 12; Gaps 2;

QY 14 TTCCAATCTTTAGATTCGATGCGCACTTAACCATTTGTCAGAGTCTGTGTGTGTGTAG 73
DB 218 TCCCAAGCTCTCGTTGATGAGAAACATGATTTAATCCGATGCTGTGTGTGTGTG 277
QY 74 CTGAAGCTTTATCAAGGTTTGTGCTCAAGTTCGATTCGATTCCTTGTGTGTGTG 133
DB 278 CTGCGGTGTATCAAGAGCTTTATTAAGCAAGGTCGATGTTCATTAAGCATATGA 337
QY 134 ACCAAGAAAGATCTGTGCAACAGCGAATACCAAAATAGCTACTGAAATTTGA 193
DB 338 ACTTGAAAGAACACACACAGAGCTGTAGAGCTTTCAATGCGGCGAAGAGCAATGA 397
QY 194 AGTGAAGAGATTCCTCA-----AGATGGTTCAATGCTGTGTATTTGTGATTT 244
DB 398 AAGGTAATTCGAATTCACCAATCGGTCAAGTGAAGTCTGTGTATTAATTTGCGATG 457
QY 245 CTGATACCGTTCAAGGTTTGTCTCAAGTTGTCAAGATTTTGTGAAGTTCCTATTC 304
DB 458 CTGAAGCTGTGCACTTGAATTCACCAAGCATCAAGAACACAGGCAAAATCTCAAGT 517
QY 305 ACTGTGTATCAACAGCTGTGTATCTGTGAATTTCCCATGTGAAGATTTACCAAGCA 364
DB 518 TCTTGTGTCAACATCTGCGTTACGCTGAATCTTCCAGCTGAAAGATGACCAAGCA 577
QY 365 ACGCTGAGAGATGTGAAGGTTTACTTGTGGGTTCTTTGTATTTTCTCAAGCTTTG 424
DB 578 ACGCTGAAGATCTTATGAAGTTTACGGGTTGAGGTTCACTTCTTCCCAAGCTTTTG 637
QY 425 CTAAAGCATTTATCAAGAGATTTCAAGGTTCTTGTGTGTGTGTGTGTGTGTGTGT 484
DB 638 CTAAAGCATTTATCAAGAGATTTCAAGGTTCTTGTGTGTGTGTGTGTGTGTGTGT 694
QY 485 CTGT 544

DB 695 CCGGTACCATCTGTCAACAGCCACCAACCAATGATGTACAAATGTCCAAAGCCGGTG 754
QY 545 TTATCATTTGAGCTTAAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 604
DB 755 TCAATCATTTTACCAAGATCATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 814
QY 605 CTTTAAACCCAGGTTTACATCT 625
DB 815 CATGTGTCCCGGACATCT 835

RESULT 4
US-09-248-796A-2490

/ Sequence 2490, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ NUMBER OF SEQ ID NOS: 28208
/ LENGTH: 927
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-2490

Query Match 10.6%; Score 86.6; DB 4; Length 927;
Best Local Similarity 52.0%; Pred. No. 8.1e-17;
Matches 328; Conservative 0; Mismatches 279; Indels 24; Gaps 5;

QY 182 CTGAAGATTTGAAGTTGAAGAAAGTTCCAAAGATGGGTTCAATGCTGTGTATTTCTG 241
DB 308 CTGAATATTTAACAAGCTTTATGGGTTAATCCAAAGCATCAAAATGATATTTACTG 367
QY 242 ATTGTATACCGTTCAAGGTTTGTCTCAAGTTGTCAAGATTTTGTGAAGTTCAT 301
DB 368 ATTCCAAAGATTTGAAGAAAGTTGTCAACAAATGAATGATTCGATTCATTTGATA 427
QY 302 TGCATTTGTATCAACAGCTGTGTACTG--TGAAATCTTCCATGTGAATTTACCAG 358
DB 428 TCTTGTGTCCAAAGT 487
QY 359 CCAAGAACGCTGAGAAAGTGTGAAGGTTTACTTGTGTGTGTGTGTGTGTGTGTGTGT 418
DB 488 TCGAACAATGAAACAAAGTTTGTATGTGAATTTTAAAGAGTTTATTTATGTCTCATG 547
QY 419 CTTTGTCAAGCATTTATCAAGAGATTTCAAGGTTTGTGTGTGTGTGTGTGTGTGT 478
DB 548 TTTGTGTCAATTTTCAAGAAAGGTTA-----AAGGTTCAATTTTCACTGCA 601
QY 479 CTATGTCTGT 538
DB 602 GTATGT 661
QY 539 CTGTGTGTATTCATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 598
DB 662 CTGGGTGTCAACATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 718
QY 599 TTAATTTTAAACCCAGGTTTATCAAGGTTCTTGTGTGTGTGTGTGTGTGTGTGT 658
DB 719 TCAATTTGT 769
QY 659 ACGAAGATTTATCAAGATGTATCTGTGTATCCCAACAAAGATGTCCAAACCA 718
DB 770 ATCCGAGTGTCAAGATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 829

NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 2488
LENGTH: 336
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-2488

Query Match
Best Local Similarity 52.9%; Score 55; DB 4; Length 336;
Matches 171; Conservative 0; Mismatches 140; Indels 12; Gaps 2;

QY 491 CCATTGTCAACGATCTCTCAAAACCAAGTGTCTACAAACATGTCACAGGCTGGTATCC 550
DB 20 CCCGGGTGAATACCTCAATTACAGCTCTTATTAATCTGTAAAGCTGATGACTC 79
QY 551 ATTGGCTTAAGCTTTGGCTTGTGAATGGCTTAAGTCAACATCAGATTAATCTTTAA 610
DB 80 ATTATGCCAATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 139
QY 611 ACCGAGTTACATCTACGCTCTCTTGAACCAAGATTTATCATGTATGTAACGAAGATTGT 670
DB 140 CTCACGGGTATATTA-----TTGACTGATATTTGCTGATTCAGAAATGA 190
QY 671 ACAACAGATGATCTCTGTATCCCAACAAGATGTCGACCAACCAAGAAATCATTTG 730
DB 191 AAAAAGATGTGGCAATTGACACCTTTGGAGAGAGAGATTACCAAGAAATTAATG 250
QY 731 GTGCTTTTGTACTTGTCTTGAATCTGTCTTCACTACATCTGCTGCTGCTGCTGCTGCT 790
DB 251 GGGCATATTATTAATCTTGGCTC---AAATGATCAACTATATCTATGTTCAAAATTTG 307
QY 791 TGGTGAATGCTGCTTCACTTCT 813
DB 308 CTGTTGATGGGGTATATCATGT 330

RESULT 8

US-09-248-796A-6987/C
Sequence 6987, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 6987
LENGTH: 372
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-6987

Query Match
Best Local Similarity 62.4%; Score 53; DB 4; Length 372;
Matches 83; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 14 TTCGAATTGATTCGATGCGCACTTAACCATGTGACAGTGCCTGTGGTGTAG 73
DB 205 TCCCAAGCTTCGGTTGATGAGAAACATGATATTAACCGGTGCTCTGTGGTTGG 146
QY 74 CTGAAGCTTAACTAAGGCTTTGTTGGCTTACGCTTCACTGATGCTTTGATATCG 133
DB 145 CTGCGGTGATCAAGACCTTATTAAGCAAGGTCGAGATGTGCATTAATGATGA 86
QY 134 ACCAAGAAAGAC 146
DB 85 ACTTGAAAGAAC 73

RESULT 9
US-09-134-001C-725
Sequence 725, Application US/09134001C
Patent No. 6180370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 725
LENGTH: 774
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-725

Query Match
Best Local Similarity 47.8%; Score 53; DB 3; Length 774;
Matches 223; Conservative 0; Mismatches 235; Indels 9; Gaps 2;

QY 200 AAGAAGTCCAAAGATGGCTCATATGCTGTGATTTCTGATCTGATACCTTCACA 259
DB 182 AATCAGGTGTTAAACCAATCATTTAAATTAAGATTAACATCTAATGAAGTAATC 241
QY 260 AGTGTGCTCAAGTGTCTAAGATTTTGGTAAAGTCCATTTGCACTTGTGTAACAG 319
DB 242 AATTTGTTATATGATTAAGGAATATGAAATATGATATTC---TTGTCATTAACG 298
QY 320 CTGTTACTGTGAACACTTCCCATGTGAAGATTAACCCAGCAAGCCGTGAAGATGG 379
DB 299 CTAGATCTCAATTCAGATGATCTGAATAATTTTGTATGAAGATGTTAAAGAA 358
QY 380 TGAAGTTAACTTTGGTCTTTTATGTTTCTCAAGCCTTGTCTAAGCATGATCA 439
DB 359 TCAATTAAGCTTAAATGAGCAATCTCTGTAGGCAACAGTGAAGTCAAAATGATG 418
QY 440 AAGAAGTATCAAGGCTCTGTGTTTGTGATTTGTTATGCTGTATGCTGTGCAATGCA 499
DB 419 AGAAGGCA-----GTGTTGATGATTAATGCTCTTCACTTGAATGATGATGCTA 472
QY 500 ACGATCTCAAAACCAAGTGTCTAACAATGTCCAAGCTGTGTATTCATTTGGCTA 559
DB 473 ATAAAGCAAGACCAAGCTCTTATGAAGCTTAAAGCCGGTGTGACATGCTGACTA 532
QY 560 AGACTTGGCTTGTGATGGCTTAAGTCAACATCAGATTAATTTTAAACCAAGTT 619
DB 533 AAGCTTAGCTAAGAAATGTGCTAATATGAATTAAGTAAAGCAATTCGCCCTGTT 592
QY 620 ACATCTACGTCCTTGTGACCAAGATTTATCATGATGTAACGAAGAA 666
DB 593 ACATGAACAATCAAGAAACGAAAGATTAATTAATGATTAATCTGAA 639

RESULT 10

US-09-248-796A-2730
Sequence 2730, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725

QY 14 TTCGAATTGATTCGATGCGCACTTAACCATGTGACAGTGCCTGTGGTGTAG 73
DB 205 TCCCAAGCTTCGGTTGATGAGAAACATGATATTAACCGGTGCTCTGTGGTTGG 146
QY 74 CTGAAGCTTAACTAAGGCTTTGTTGGCTTACGCTTCACTGATGCTTTGATATCG 133
DB 145 CTGCGGTGATCAAGACCTTATTAAGCAAGGTCGAGATGTGCATTAATGATGA 86
QY 134 ACCAAGAAAGAC 146
DB 85 ACTTGAAAGAAC 73

PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 2730
LENGTH: 825
TYPE: DNA
ORGANISM: *Candida albicans*
US-09-248-796A-2730

Query Match 6.3%; Score 51.8; DB 4; Length 825;
Best Local Similarity 45.9%; Pred. No. 7.3e-06;
Matches 299; Conservative 0; Mismatches 337; Indels 15; Gaps 3;

```
QY 22 TTTGATTCGATGGCCACTTAACCATTCGACAGTGGCTGGTGGTTAGTCGAAGCT 81
DB 25 TTCAATGTTAAATGGAAATGCTGTGTAACTGCTGCTGCTGCTGCTGCTGCTTATAC 84
QY 82 TTAATCAAGGGTGGTGGCTGACGCTTCTGACATTCGCTTGTGATTCGACCAAGAA 141
DB 85 TGTGCTGAGGCTTTGTTATTTGAACGGTCCCTCACTGCTGCTATTAACATCAAGAAAGCA 144
QY 142 AAGACTGCTGCAACCAAGCCGAATACCAAAATACGCTACTGAAGATTGAAGTTGAA 201
DB 145 AAGCTTGTGAAGAGCTCAAAAGTATTTGGAAAAAATTGCCAAAGACACACAGAA- 203
QY 202 GAAGTCCAAAGATGGCTTCAATGCTGATGATATTTCTGATTCGATACGCTTCAAG 261
DB 204 -----TTGTAAATTAATTTCAATTCCTGCAATATGCTGTTGAAGAAATGTGAANA 258
QY 262 GTGTTTGTCAAGTTGCTTAAGATTTTGTGAATTTGATTCGCTTGTGCTTGAACAGCT 321
DB 259 TTCTAGCCGCAAAATAGCAACAACTGCAAAAGTTGACAT---TTTGTGGCCAAAGCT 315
QY 322 GGTACTGTGAAAACCTTCCATGTAAGATTAACCAAGCAAGAGCTGGAAGATGCTG 381
DB 316 GGTGCTAGTGGGGTGTCCGTTGAGAAAGCAATCAAGTTTGAAGCAGTCAAAAAGTCTC 375
QY 382 AAGTTACTGTTGGGTTCTTGTGATTTTCTCAAGCTTTGCAACCAATGATCAAA 441
DB 376 AACTGGAATGTTGTTGCGGTCAACACACATCAAGTTATTTACCTTTATTTGAAAAA 435
QY 442 GAAGTATCAAGGGTGTCTTGTGTTTGTGATTTGT-----TCTATGCTGCTGCCATT 495
DB 436 GCTGTATCAAAAGAGATCTTCTGATTTGATTCCTTATGCTCTGCTGATTAAGTTATCG 495
QY 496 GTCAACGATTCCTCAAAACCAAGTTGTCTACCAACATGTCACAGCTGGTGTATTCATTG 555
DB 496 ACTACGATATGTCGCTGTATACGCTTATTTGCTTCCAGAGGCTGCAGTATCTCATTTG 555
QY 556 GCTAAGACTTTGGCTTGTGAATGGGCTAAGTACATCAAGGTTATTTCTTAACCA 615
DB 556 GGAAGAAACTGTGTCTCAATTTGCCCCAGACATCAATGTTAATTCATGCCCA 615
QY 616 GGTTCATCTACGCTTCCTTGTGACCAAGATTTATCAATGTAAGCAAGAA 666
DB 616 GGTTTTTCCTCATCAAAATGSCCAATGTTGATGCAAGCTGCTGTGA 666
```

RESULT 11
US-09-367-012-2

Sequence 2, Application US/09367012

Patent No. 6218156

GENERAL INFORMATION:

APPLICANT: Yasohara, Yoshihiko

APPLICANT: Kizaki, No. 62181561yuki

APPLICANT: Hasegawa, Junzo

APPLICANT: Wada, Masaru

APPLICANT: Shimizu, Sakayu

APPLICANT: Katoaka, Michihiko

APPLICANT: Yamamoto, Kazuhiko

APPLICANT: Kawabata, Hiroshi

APPLICANT: Kita, Keiko

TITLE OF INVENTION: No. 62181561 carbonyl reductase, gene coding same, and method
TITLE OF INVENTION: for using such reductase and gene
FILE REFERENCE: S72030PCUS
CURRENT APPLICATION NUMBER: US/09/367,012
CURRENT FILING DATE: 1999-11-24
EARLIER APPLICATION NUMBER: PCT/J97/03051
EARLIER FILING DATE: 1997-09-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 852
TYPE: DNA
ORGANISM: *Candida magnoliae*
US-09-367-012-2

Query Match 6.3%; Score 51.2; DB 3; Length 852;
Best Local Similarity 58.6%; Pred. No. 1.2e-05;
Matches 89; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```
QY 437 TCAAGAGGATATCAAGGCTGCTTGTGTTTGTGATTTGATTCATGCTGTCGCTTGG 496
DB 485 TCGAAGAGAGGGCAAGAGGGCGCTTGTGTTCAAGGCTTCATGCTGCGCCACATTG 544
QY 497 TCAAGATCTCTCAAAACCAAGTTGTCTACAAATGTCAGAGCTGCTGTTATCCATTGG 556
DB 545 TGAACGTGCCCCAGTTTCAGGCCACGTACAAACGGCGCCAGGCTGGCGCCACATTG 604
QY 557 CTAGACTTTGGCTTGTGAATGGGCTAAGTAC 588
DB 605 CGAAGTCTGGCCGCTGAGTTGGCGCGCTTC 636
```

RESULT 12
US-09-777-157A-2

Sequence 2, Application US/09777157A

Patent No. 6448052

GENERAL INFORMATION:

APPLICANT: Yasohara, Yoshihiko

APPLICANT: Kizaki, No. 64480521yuki

APPLICANT: Hasegawa, Junzo

APPLICANT: Wada, Masaru

APPLICANT: Shimizu, Sakayu

APPLICANT: Katoaka, Michihiko

APPLICANT: Yamamoto, Kazuhiko

APPLICANT: Kawabata, Hiroshi

APPLICANT: Kita, Keiko

TITLE OF INVENTION: Carbonyl Reductase Enzyme and Methods for its Use

FILE REFERENCE: 068383.0110

CURRENT APPLICATION NUMBER: US/09/777,157A

CURRENT FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: US 09/367,012

PRIOR FILING DATE: 1999-11-24

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentin version 3.0

SEQ ID NO 2

LENGTH: 852

TYPE: DNA

ORGANISM: *Candida magnoliae*
US-09-777-157A-2

Query Match 6.3%; Score 51.2; DB 3; Length 852;
Best Local Similarity 58.6%; Pred. No. 1.2e-05;
Matches 89; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```
QY 437 TCAAGAGGATATCAAGGCTGCTTGTGTTTGTGATTTGATTCATGCTGTCGCTTGG 496
DB 485 TCGAAGAGAGGGCAAGAGGGCGCTTGTGTTCAAGGCTTCATGCTGCGCCACATTG 544
QY 497 TCAAGATCTCTCAAAACCAAGTTGTCTACAAATGTCAGAGCTGCTGTTATCCATTGG 556
DB 545 TGAACGTGCCCCAGTTTCAGGCCACGTACAAACGGCGCCAGGCTGGCGCCACATTG 604
QY 557 CTAGACTTTGGCTTGTGAATGGGCTAAGTAC 588
```

Db 605 CGAAGTCGCTGCGCTCGAGTTCGCGCGCTTC 636

RESULT 13
US-09-734-237B-69

; Sequence 69, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent version 3.1
; SEQ ID NO 69
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-734-237B-69

Query Match 6.3%; Score 51.2; DB 4; Length 852;
Best Local Similarity 58.6%; Pred. No. 1.2e-05;
Matches 89; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 437 TCAAGAAGATATCAAGGCTGCTTTGTTTATTTGTTATGTTGTCGTCATTTG 496
Db 485 TCGAGAAAGAGGCGCAAGAGGCGCTTGTGTTACGCGCTTCATGCTGCGCCACATTTG 544
Qy 497 TCAAGATCTCTCAAAACCAAGTTGTCTACAACATGTCACAGGCTGTGTTATTCATTTGG 556
Db 545 TGAAGTGTCCCAAGTTCAGGCGCACGTCACACGCGGCGCAAGGCTGCGCTGCGCCACTTGG 604
Qy 557 CTAAGACTTTGGCTTGTGATGAGGCTAAGTAC 588
Db 605 CGAAGTCGCTGCGCTCGAGTTCGCGCGCTTC 636

RESULT 14
US-09-489-039A-5819

; Sequence 5819, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5819
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5819

Query Match 6.3%; Score 51.2; DB 4; Length 951;
Best Local Similarity 54.9%; Pred. No. 1.2e-05;
Matches 101; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 454 GGTGCTCTGTGTTTATTTGTTATGTTGTCGTCATGTCACAGATCTCAAAAC 513
Db 595 GCGGAGTCAATTAATCAATATGCGCTCGATGTCGGGATTAATTTAAACGCGGCGCTCGAC 654
Qy 514 CAAATTGCTACACATATCCAGGCTGTGTATCCATTTGGCTAAGACTTTGGCTTGT 573

Db 655 CAGGCGCATTTACACTGCTCCAAAGCGCGGTATCATCTGTCCAAAAAGCTGCGCATG 714
Qy 574 GAATGGGCTAAGTACACATGACAGTAAATTTCTTAAACCGAGTTACATTCAGATCCT 633
Db 715 GAGTGGGTGGCGAAAGGATCGCGTCAATTCATCAGTCCGGGATATACCGCCACGCCG 774
Qy 634 TTGA 637
Db 775 ATGA 778

RESULT 15
US-09-367-012-8

; Sequence 8, Application US/09367012
; Patent No. 6218156
; GENERAL INFORMATION:
; APPLICANT: Yasohara, Yoshihiko
; APPLICANT: Kizaki, No. 6218156iyuki
; APPLICANT: Hasegawa, Junzo
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Kataoka, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; APPLICANT: Kita, Keiko
; TITLE OF INVENTION: No. 6218156el carbonyl reductase, gene coding same, and method
; FILE REFERENCE: S72030PUS
; CURRENT APPLICATION NUMBER: US/09/367,012
; CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: PCT/JP97/03051
; EARLIER FILING DATE: 1997-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1303
; TYPE: DNA
; ORGANISM: Candida magnoliae
US-09-367-012-8

Query Match 6.3%; Score 51.2; DB 3; Length 1303;
Best Local Similarity 58.6%; Pred. No. 1.5e-05;
Matches 89; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 437 TCAAGAAGATATCAAGGCTGCTTTGTTTATTTGTTATGTTGTCGTCATTTG 496
Db 643 TCGAGAAAGAGGCGCAAGAGGCGCGCTTGTGTTACGCGCTTCATGTTGCGCCACATTTG 702
Qy 497 TCAAGATCTCTCAAAACCAAGTTGTCTACAACATGTCACAGGCTGTGTTATTCATTTGG 556
Db 703 TGAAGTGTCCCAAGTTCAGGCGCACGTCACACGCGGCGCAAGGCTGCGCTGCGCCACTTGG 762
Qy 557 CTAAGACTTTGGCTTGTGATGAGGCTAAGTAC 588
Db 763 CGAAGTCGCTGCGCTCGAGTTCGCGCGCTTC 794

Search completed: July 12, 2005, 15:30:36
Job time : 191 secs

This Page Blank (uspio)

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|---------------------|-----------------------|
| 1 | 81.6 | 100.0 | 816 | 19 | US-10-720-018-1 | Sequence 1, April |
| 2 | 99.4 | 12.2 | 3203 | 20 | US-10-653-047-61 | Sequence 41, April |
| 3 | 88.6 | 10.9 | 687 | 20 | US-10-653-047-4932 | Sequence 4932, April |
| 4 | 81.8 | 10.0 | 637 | 17 | US-10-369-493-36723 | Sequence 36723, April |
| 5 | 80.6 | 9.9 | 735 | 17 | US-10-369-493-26991 | Sequence 26991, April |
| 6 | 74.6 | 9.1 | 852 | 10 | US-09-734-227B-71 | Sequence 71, April |
| 7 | 74.2 | 9.1 | 146 | 20 | US-10-653-047-4825 | Sequence 4825, April |

61 100Y

```
Db 61 TGTGTGTTTAAAGTGAAGCTTATCAAGGTTTGTGGCTACGGTTCGACATTGCT 120
Qy 121 TTGCTTATATCGACCAAGAAAGAGCTGCTCCCAACAGCCGAATACCAATAGCT 180
Db 121 TTGCTTATATCGACCAAGAAAGAGCTGCTCCCAACAGCCGAATACCAATAGCT 180
Qy 181 ACTGAAGATTTGAAGTTGAAGAGTTCCCAAGATGGGTTCAATGCTGATATTTCT 240
Db 181 ACTGAAGATTTGAAGTTGAAGAGTTCCCAAGATGGGTTCAATGCTGATATTTCT 240
Qy 241 GATTCTGATACCTGTCACAGAGTGTGCTCAAGTTCGTAAGATTTTGTAGTGGCA 300
Db 241 GATTCTGATACCTGTCACAGAGTGTGCTCAAGTTCGTAAGATTTTGTAGTGGCA 300
Qy 301 TTGCACTGTTTAAACAGCTGCTGTTACCTGTAAGAACTCCATGTAAGATTAACCCAGCC 360
Db 301 TTGCACTGTTTAAACAGCTGCTGTTACCTGTAAGAACTCCATGTAAGATTAACCCAGCC 360
Qy 361 AAGAAGCTGAGAGATGGTGAAGGTTAACTGTTGGGTTCTTGATGTTTCTCAAGCC 420
Db 361 AAGAAGCTGAGAGATGGTGAAGGTTAACTGTTGGGTTCTTGATGTTTCTCAAGCC 420
Qy 421 TTGCTTAAGCCATTGATCAAGAGATCAAGAGGTTGCTTGTGTTGATGTTGTTCT 480
Db 421 TTGCTTAAGCCATTGATCAAGAGATCAAGAGGTTGCTTGTGTTGATGTTGTTCT 480
Qy 481 ATGTCTGCTGCTTGTCAAGATCTCTCAAAACCAAGTTGTCACAAATGTCAGAGCT 540
Db 481 ATGTCTGCTGCTTGTCAAGATCTCTCAAAACCAAGTTGTCACAAATGTCAGAGCT 540
Qy 541 GGTGTATTCATTTGGCTTAAGACTTGGCTTGAATGGGCTTAAGTCAACATCAGAGTT 600
Db 541 GGTGTATTCATTTGGCTTAAGACTTGGCTTGAATGGGCTTAAGTCAACATCAGAGTT 600
Qy 601 AATCTTTAAACCCAGTTACATCTACGCTCTTTGACCAAGATGTTATCAATGTAAC 660
Db 601 AATCTTTAAACCCAGTTACATCTACGCTCTTTGACCAAGATGTTATCAATGTAAC 660
Qy 661 GAAGATTTGACACAGATGATCTCTGATATCCCAACAAAGATGCCAACAAG 720
Db 661 GAAGATTTGACACAGATGATCTCTGATATCCCAACAAAGATGCCAACAAG 720
Qy 721 GAATCATTTGGTCTGTTTGTACTGCTTCTGAATGCTGCTTCACTACTGCT 780
Db 721 GAATCATTTGGTCTGTTTGTACTGCTTCTGAATGCTGCTTCACTACTGCT 780
Qy 781 GCCAGCTTACTGTTGATGATGTTGCTTCACTTCTGG 816
Db 781 GCCAGCTTACTGTTGATGATGTTGCTTCACTTCTGG 816

RESULT 2
US-10-653-047-41
; Sequence 41, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groch Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
```

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 3203
; TYPE: DNA
; ORGANISM: Fusarium venenatum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (3203)
; OTHER INFORMATION: n = A,T,C or G
US-10-653-047-41

Query Match 12.2%; Score 99.4; DB 20; Length 3203;
Best Local Similarity 52.7%; Pred. No. 4.5e-18;
Matches 265; Conservative 0; Mismatches 211; Indels 7; Gaps 2;

Qy 308 TGGTTACACAGCTGGTTACTGTGAAGAACTTCCATGTAAGATTAACCAAGCAAGACG 367
Db 2614 TAGTCACTGGCTGGCTGCTTACCGAGAAATTTCCAGGCAATCACTATCCATGATGTA 2673
Qy 368 CTGAGAAATGCTGAAGGTTAACTGTTGGGTTCTTGATGTTTCTCAAGCTTGTCTA 427
Db 2674 TGGCAAGTTGTGGGGTGTAAATGTGACGGTACTTCTTGTGCAAGTTGCAAGTTCCA 2733
Qy 428 AGCATTGATCAAGAGATCAAGAGTCTCTGTTGTTGATGTTGATGTTGATGTTGAT 487
Db 2734 A--ACATCTCATGAAAGCCAGGTAACCTGTAATGATGATGATGATGATGATGATG 2790
Qy 488 GTGCTATGTCACAGATCTCAAAACCAAGTTGTCACAAATGTCAGAGCTGTTGTTA 547
Db 2791 GTGCTATGTCACAGATCTCAAAACCAAGTTGTCACAAATGTCAGAGCTGTTGTTA 547
Qy 548 TCCATTTGGCTTAAGACTTGGCTTGAATGGGCTTAAGTCAACATCAGATTAATTTCT 607
Db 2851 GTCACTGGCTGCTTCCCTCGCAGTGAAGTGGCTCAAGCTGTAATCCAGTCACTGTA 2910
Qy 608 TAAACCAAGTTACATCTAGAGTCTTGTGACCAAGATGTTATGATGTAAGAGAT 667
Db 2911 TCTCTCTGGCTTAAGTGTGATCTCTTAAGCAAGAGATTTGAAGATTAACCAAGTA 2970
Qy 668 TGTCAACAGATGATCTGTAATCCCAACAAAGATGCCAACAAGATTAACCAAGATTA 727
Db 2971 TCGAGAGACCTGACATCTTATCTCAGGAGATGATGATGATGATGATGATGATGAT 3030
Qy 728 TTGCTGTTTGTGATCTTCTGATGATGCTGCTTCACTACTACTGCTGCTGCT 787
Db 3031 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3086
Qy 788 TACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810
Db 3087 TTAGAGTAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3109

RESULT 3
US-10-653-047-4932
; Sequence 4932, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groch Clausen
; APPLICANT: Peter Bjørke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
; CURRENT APPLICATION NUMBER: US/10/653,047
; PRIOR FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/533,559
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/273,623
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 7860
```

SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 4932

LENGTH: 687
 TYPE: DNA
 ORGANISM: Aspergillus oryzae
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)...(687)
 OTHER INFORMATION: n = A,T,C or G
 US-10-653-047-4932

Query Match 10.9%; Score 88.6; DB 20; Length 687;
 Best Local Similarity 49.8%; Pred. No. 2.9e-15;
 Matches 279; Conservative 0; Mismatches 275; Indels 6; Gaps 2;

196 TTGAAGAAGTTCAGAAAGATGGCTCATATGCTGTGATTTCTGATTTCTGATACCGTT 255
 114 TTGAAGCAAAATGCCCAAGCTACCGCCACTAGCTGATTTCCGACCTTAACCTCCGTC 173
 256 CACAAAGTGTTCCTCAAGTTCCTAAGATTTTGTGAAGTTCCTTCACTTGTGTTAAC 315
 174 AACGATGCCCTCTCCGATTTATCTCAAGCAGCGCAAGAT---GACAACTGGTCAAC 230
 316 ACAGTGTGTTACTGTGAAGAACTTCCCATGTGAAGATTAACCCAGCCAGAACGCTGAGAG 375
 231 TCCGCGGATTCACGAAAACTTGATGCAATCTCTTACCTTCACAGCCGCTGCAAAAG 290
 376 ATGTGAAGTTCATCTGTGGCTCTTTGTATGTTTCTCAAGCTTGTCTAAGCCATTG 435
 231 CTTGGGGCGTTAATGTGATGGAACATACCTTTTCGCAACCGGTGTGCCAAGC---AC 347
 436 ATCAAGAAGGTATCAAGGGTCTTCTGTGTTTGTATGTTGTTCTATGTCTGTGCCANT 495
 348 CTCAATGAGCGCAAGGTTCCGGGAGAGATTTGATGATGATGATGATGATGATGATGATG 407
 436 GTCAACGATCTCTCAAAACCAAGTGTCTTCAACATGTCCAAAGGCTGTGTATTCATTG 555
 408 GACACAGCGCGCAGCCAGCCAGCTCTTCAACAGCGCGCGCTGTGTCTCAACTT 467
 556 GCTAAGACTTTGGCTGTGAATGGGCTAAGTACATCAAGATTAATCTTTAAACCA 615
 468 GCCGGTACTTCCCGCGGAATGGCCGGTACGACATCCGGGTAGCTGATGAGCCCT 527
 616 GGTATCATCTACGGTCTTTTGAACCAAGATTTTCAATGTGAACGAAGATTTGATCAAC 675
 528 CGATTCATGCTTACTGCTCCCTGACCGCAAGATTTTGAATGAAGACCCCGATTTGGGGAC 587
 676 AGATGATCTCTGTGATTCACCAAGATGTCCGAACCAAGGAATATCTGTGCT 735
 588 AAGTGAATCTGCTCATCCCAACCGGCAAGATGGGTACTCCGAGGACCTGATGGTCCC 647
 736 GTTTGTACTGCTTTCTGA 755
 648 GTTACCTTTCTGCTCAGGA 667

RESULT 4

US-10-369-493-36723
 Sequence 36723, Application US/10369493
 Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 FILE REFERENCE: 38-10(52052)B
 CURRENT APPLICATION NUMBER: US/10/369,493
 CURRENT FILING DATE: 2003-02-28
 PRIOR APPLICATION NUMBER: US 60/360,039
 PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374
 SEQ ID NO 36723
 LENGTH: 637
 TYPE: DNA
 ORGANISM: Aspergillus nidulans
 US-10-369-493-36723

Query Match 10.0%; Score 81.8; DB 17; Length 637;
 Best Local Similarity 55.0%; Pred. No. 2.8e-13;
 Matches 183; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

304 CACTTGCTTACACAGCTGTGTACTGTGAAAACTTCCATGTGAAGATTAACCCAGC 363
 70 CATCTGTACGCTCGCGGATTTACGAAACCTTGACGCCCATGACCTACCAACAGC 129
 364 AACCTGGAAGATGTGAAGATTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 423
 130 CGTATGCAAGACTTGGGGGTGAACCTGACGCAATATCTTTGCTACGGGGCTG 189
 424 GCTAAGCATTTGATCAAGAAAGTATCAAGGGTCTTCTGTGTTTGTATGTTGTTGTTGTT 483
 190 GCAAGCACTTGAT---GAGGCGCAAGCGCCCGGAGATGATGATGATGATGATGATGATG 246
 484 TGTGTGCTATTTGATCAAGATTTCTTCAAAACCAAGTTGTCTTACATGTCAGAGCTGT 543
 247 TCTGTGCTATCTGTCAAGTCTTCAACCAAGGCGCTTCAACGCGCCAGGAGCAGC 306
 544 GTTATCCATTTGGCTTGAAGCTTTGGCTTGTGAATGGGCTAAGTACATCAAGATTAAT 603
 307 GTTGCACCTTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 366
 604 TCTTAAACCAAGGTTTACATCTACGCTTTG 636
 367 TGCATTAACCTCGATACATGCTGACTGCTG 399

RESULT 5

US-10-369-493-26991
 Sequence 26991, Application US/10369493
 Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 FILE REFERENCE: 38-10(52052)B
 CURRENT APPLICATION NUMBER: US/10/369,493
 CURRENT FILING DATE: 2003-02-28
 PRIOR APPLICATION NUMBER: US 60/360,039
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 47374

US-10-369-493-26991
 Sequence 26991, Application US/10369493
 Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Slater, Steven C.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 FILE REFERENCE: 38-10(52052)B
 CURRENT APPLICATION NUMBER: US/10/369,493
 CURRENT FILING DATE: 2003-02-28
 PRIOR APPLICATION NUMBER: US 60/360,039
 PRIOR FILING DATE: 2002-02-21
 NUMBER OF SEQ ID NOS: 47374

304 CACTTGCTTACACAGCTGTGTACTGTGAAAACTTCCATGTGAAGATTAACCCAGC 363
 70 CATCTGTACGCTCGCGGATTTACGAAACCTTGACGCCCATGACCTACCAACAGC 129
 364 AACCTGGAAGATGTGAAGATTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 423
 130 CGTATGCAAGACTTGGGGGTGAACCTGACGCAATATCTTTGCTACGGGGCTG 189
 424 GCTAAGCATTTGATCAAGAAAGTATCAAGGGTCTTCTGTGTTTGTATGTTGTTGTTGTT 483
 190 GCAAGCACTTGAT---GAGGCGCAAGCGCCCGGAGATGATGATGATGATGATGATGATG 246
 484 TGTGTGCTATTTGATCAAGATTTCTTCAAAACCAAGTTGTCTTACATGTCAGAGCTGT 543
 247 TCTGTGCTATCTGTCAAGTCTTCAACCAAGGCGCTTCAACGCGCCAGGAGCAGC 306
 544 GTTATCCATTTGGCTTGAAGCTTTGGCTTGTGAATGGGCTAAGTACATCAAGATTAAT 603
 307 GTTGCACCTTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 366
 604 TCTTAAACCAAGGTTTACATCTACGCTTTG 636
 367 TGCATTAACCTCGATACATGCTGACTGCTG 399

ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_88858C.1
US-10-425-115-172450

Query Match 8.3%; Score 72.4; DB 20; Length 775;
Best Local Similarity 56.2%; Pred. No. 1.9e-10;
Matches 136; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 383 AGGTAACTTGTGGGTTCTTTGATGTTTCAAGCCCTTCTAAGCCATTGATCAAG 442
DB 465 AGTCAAGCTTTGGGGAAGATTGGCTGTGCTCAAGCCGCTCCAGCTTGTGACCACT 524
QY 443 AAGGATCAAGGCTGCTTCTGTTGTTTGAATGTTTCAATGTTGAGTCAAGC 502
DB 525 CTGGCTTCAAGCAGCGTCTTCATGCTTCATCCTCGCCATGCTTCCCAATGCTCAAC 584
QY 503 ATCCCAAAACCAAGTTGCTCAACATGTCACAGGCTGCTTATCCATTGGCTAAGA 562
DB 585 GTGGCATCCACCAAGCTTCTTCAACATGCTCCAGGCTGCGCTGCTGATGTTAAGC 644
QY 563 CTTCGCTTGTGAATGGGCTAAGTCAACATGATGATTAATCTTAAACCAAGTTACA 622
DB 645 AGCTGCTGTTAGTGGGCGGCAAGGGATTCGAGTCAAGCTCTTGGCCAGGTTACG 704
QY 623 TC 624
DB 705 TC 706

RESULT 9

US-10-369-493-25810
Sequence 25810, Application US/10369493
Publication No. US2003023675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 25810
LENGTH: 783
TYPE: DNA
ORGANISM: Schizosaccharomyces pombe
US-10-369-493-25810

Query Match 8.3%; Score 67.6; DB 17; Length 783;
Best Local Similarity 51.2%; Pred. No. 5.1e-09;
Matches 213; Conservative 0; Mismatches 194; Indels 9; Gaps 2;

QY 171 CAATACCTACTGAAGATTAAGTGAAGAAATTCACCAAGATGGTCTATATGCTG 230
DB 171 CGAAAGGCTCGAAGAAATGGCCAGCTAATGCTGCAAGCTTATACATGCAATG 230
QY 231 TGATATTTCTGATTTGATACCGTTCAAGGTTGTTGCTCAAGTTGTAAGATTTGG 290
DB 231 CGATCTACTATTTCTTAAAGAGTGAAGATGCTTGTGATGATTCAAAGGCTTTGA 290
QY 291 TAACTGTCATGCACTTGTGTTAACAAGCTGCTTACTGTGAAAACCTTCCATGTAAGA 350
DB 291 TACTATGATATGTTGTGCGCAACAAGGTAATTTGCACTGGCAAGTCGGCAT---GA 347
QY 351 TTACCCAGCAAGAACGCTGAGAAATGTAAGTTAACTTTGTTGGTTCTTTGTATGT 410
DB 348 CATGACTTACGAAGATTTGCTTAATGAATAAAGTCAACTTGGTGTGCTTCAACGT 407

QY 411 TTCTCAAGCTTTGCTTAAGCAATTGATCAAGAAAGTATCAAGGCTTCTGTTGTTT 470
DB 408 CGCTCACAATGCTGGCCCGCATCTTCCAAAGCAAG-----CATGGCAGAGCTTGTAGC 461

QY 471 GATTGCTTATATGCTGTTGCTGCTATGCTCAAGATCTTCAAAACCAAGTTGTATACACT 530
DB 462 TACTGCTTCAATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 521

QY 531 GTCCAAGCTGCTGTTATTCATTTGGCTAAGACTTGGCTTGTGAATGGGCTAAGT 586
DB 522 TTCAGAGCCGGTGCATTTCACTATCATCAAGAGTTTGGCAATGCGCAAT 577

RESULT 10

US-10-369-493-26610
Sequence 26610, Application US/10369493
Publication No. US2003023675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 26610
LENGTH: 774
TYPE: DNA
ORGANISM: Thermotoga maritima
US-10-369-493-26610

Query Match 7.8%; Score 63.6; DB 17; Length 774;
Best Local Similarity 61.4%; Pred. No. 7.7e-08;
Matches 102; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 472 ATGGTCTATGCTGTGCTGATTCATGTCACAGATCTTCAAAACCAAGTTGTATCAACATG 531
DB 436 ATCGGTCAATGCTCGGACACATGTAACAACCTGAGAGCAGACGCTTACACACT 495
QY 532 TCAGAAGCTGTTATTCATTTGGCTTAAGACTTGTGCTTGTGAATGGGCTTAAGTAC 591
DB 496 TCGAAAGCGGTGTGATTCATCTCACAGATCTGCGCGCGAGTGGGCGCCGTAACGGA 555
QY 592 ATCAGAGTTATTTCTTAAACCCAGGTTACATCTACGCTCTTGA 637
DB 556 ATCAGAGTGAACAGCATTAAGCCCGGATACATCAAGAACACCTCTCA 601

RESULT 11

US-10-653-047-3866
Sequence 3866, Application US/10653047
Publication No. US20040229367A1

GENERAL INFORMATION:

APPLICANT: Randy M. Berka
APPLICANT: Michael W. Rev
APPLICANT: Jeffrey R. Shuster
APPLICANT: Sakari Kauppinen
APPLICANT: Ib Groth Clausen
APPLICANT: Peter Bjarke Olsen
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE REFERENCE: 5849 200-US
CURRENT APPLICATION NUMBER: US/10/653,047
CURRENT FILING DATE: 2003-08-29
PRIOR APPLICATION NUMBER: US/09/533,559
PRIOR FILING DATE: 2000-03-22

```

: PRIOR APPLICATION NUMBER: 09/273,623
: PRIOR FILING DATE: 1999-03-22
: NUMBER OF SEQ ID NOS: 7860
: SOFTWARE: FastSeq for Windows Version 4.0.
: SEQ ID NO 3866
: LENGTH: 636
: TYPE: DNA
: ORGANISM: Aspergillus niger
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(636)
: OTHER INFORMATION: n = A,T,C or G
:
: JS-10-653-047-3866

```

| | | | | |
|---------------------------|--------|--------------------|-----------|-------------|
| Query Match | 7.4%; | Score 60.6; | DB 20; | Length 636; |
| Best Local Similarity | 51.3%; | Pred. No. 5.3e-07; | | |
| Matches 194; Conservative | 0; | Mismatches 175; | Indels 9; | Gaps 2 |

| | | | |
|----|-----|---|-----|
| QY | 211 | AAGATGGGTTCAATPAGCCGTGATCAATTTCTGATTTCTGATACCGTTTCAACAAGGTTTGCT | 270 |
| Db | 200 | AAGTCAGCTGCTATPAAAGCTGATGTGCGAAATCAGAGAGATTTAGAGCGTAAATCCAG | 259 |
| QY | 271 | CAAGTGTCAAGATTTTGTAAAGTTCCATTGCACTTGGTTAAACAAGCTGGTTACTGT | 330 |
| Db | 260 | CAGATTGCTTCAGACTTTTGGCAAGCTTGATATATATGTCGTGAC--TCGGGGGTCACT | 316 |
| QY | 331 | GAAATCTTCCATGTGAAAGTTAACCAAGCCAGAACGCTGAGAAAGATGTGAAGGTTAA | 390 |
| Db | 317 | TCCAAATATGCGGCGAGAAAGACTAACCAACACGGAACATGGCGTACATCAAGAAAGTCAAT | 376 |
| QY | 391 | TGTGGGGTCTTTGTATGTTTTCTAAGCCTTGCTTAAGCCATTGATCAAGAAGGTATC | 450 |
| Db | 377 | CTAGATGGGCAATTTTAAACCGCGAAGCGGCTTTCGATCTTCAAGCAACAAG---- | 432 |
| QY | 451 | AAGGGTCTTCTGTTGTTTGTATGTGTTCTATGTCGTGACCATTGTCAACGATCTCAA | 510 |
| Db | 433 | --ACATGAAATGTATCTTCAAGGCTGTGAGTGAACAATTTGGTAAATGTGCTCAG | 490 |
| QY | 511 | AACCAAGTGTCTACACAATGTCCAAAGGCTGTGTATTCATTGGCTAAAGACTTTGGCT | 570 |
| Db | 491 | AAGCAAGTGCATACATAGCTCTCAAAAGCTGGGGTGTGTCAAAATGGCAAAATGTCTGTCT | 550 |
| QY | 571 | TGTGATGGGCTTAAGTAC | 588 |
| Db | 551 | GTTGAATGGGTGANTTC | 568 |

```

RESULT 12
US-10-314-394-1
/ Sequence 1, Application US/10314394
/ Publication No. US20030143700A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamamoto, Hiroaki
/ APPLICANT: Ueda, Momoko
/ APPLICANT: Pan, Ritsuzui
/ APPLICANT: Hamatani, Takehi
/ TITLE OF INVENTION: METHODS FOR PRODUCING OPTICALLY ACTIVE ALCOHOLS
/ FILE REFERENCE: SHZ-006
/ CURRENT APPLICATION NUMBER: US/10/314,394
/ CURRENT FILING DATE: 2002-12-06
/ PRIOR APPLICATION NUMBER: US 60/385,434
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: JP 2002-152955
/ PRIOR FILING DATE: 2002-05-27
/ PRIOR APPLICATION NUMBER: JP 2001-375041
/ PRIOR FILING DATE: 2001-12-07
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: SeqIdIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 822
/ TYPE: DNA
/ ORGANISM: Datura stramonium
US-10-314-394-1

```

| | | | | |
|---------------------------|--------|--------------------|-----------|-------------|
| Query Match | 7.3% | Score 59.8; | DB 15; | Length 822; |
| Best Local Similarity | 52.5%; | Pred. No. 1.1e-06; | | |
| Matches 158; Conservative | 0; | Mismatches 137; | Indels 6; | Gaps 1. |

```

RESULT 13
US-10-314-394-3
/ Sequence 3, Application US/10314394
/ Publication NO. US20030143700A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamamoto, Hiroaki
/ APPLICANT: Ueda, Momoko
/ APPLICANT: Pan, Ritsuzui
/ APPLICANT: Hamatani, Takeshi
/ TITLE OF INVENTION: METHODS FOR PRODUCING OPTICALLY ACTIVE ALCOHOLS
/ FILE REFERENCE: SHZ-006
/ CURRENT APPLICATION NUMBER: US/10/314,394
/ CURRENT FILING DATE: 2002-12-06
/ PRIOR APPLICATION NUMBER: US 60/385,434
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: JP 2002-152855
/ PRIOR FILING DATE: 2002-05-27
/ PRIOR APPLICATION NUMBER: JP 2001-375041
/ PRIOR FILING DATE: 2001-12-07
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 825
/ TYPE: DNA
/ ORGANISM: Hyoscyamus niger
US-10-314-394-3

```

| | Query Match | Similarity | Score | DB | length |
|----|-------------|-----------------------------|---------------------------------|-------------|---------|
| | Best Local | 52.4% | Pred. | No. 1.le-05 | 825; |
| | Matches | 151; Conservative | 0; Mismatches | 131; Indels | 6; Gaps |
| | | | | | 1 |
| Qy | 532 | TCCAAGCTGCGTATCCATTGGCTAGA | CTTTGGCTTGTAATGGGCTAAGTACAAC | 591 | |
| | | | | | |
| Dd | 523 | TCCAAGCTGCATAATCAATTACGAAGA | CTTGGCATGTGAATGGGCCAAGACAAC | 582 | |
| | | | | | |
| Qy | 592 | ATCAGAGTTAATCTTTAAACCAGGTTA | CATCTACGGTCCTTTGACCAAGATGTTATC | 651 | |
| | | | | | |
| Dd | 583 | ATTGGGTCATTCAGTCTCTCAGAGAG | CATTTTAACCCCACTATGAAACTCAATT | 642 | |
| | | | | | |
| Qy | 652 | -----AATGCTAACAGAAATTGTAACA | AGATGGATCTCGGTATCCCAACAAGAGA | 705 | |
| | | | | | |
| Dd | 643 | AAGAAAAATCCCTATCCAAAAGAAAG | AATAGCAATTTTATTGTCAAGCTCCATGGGC | 702 | |
| | | | | | |

QY 706 ATGTCCGAACAAGGATATCATGTGCTGTTTGTACTGTTCTGATCTGCT 765
 DB 703 CGGGCTGGAAAGCCCAATGAGTGTCTGACTTAATAGCCCTTCTGCTCCCTGCTGCT 762
 QY 766 TCATCACTACTGTCGTCAGCTTACTGTTGATGTTGTTTCACTTCT 813
 DB 763 TCTTATATTAATCTGCGCAATTAATATGCGCTGATGTTGATTCACAGCT 810

RESULT 14

US-09-938-842A-948
 ; Sequence 948, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 948
 ; LENGTH: 789
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana

US-09-938-842A-948

Query Match 6.9%; Score 56.2; DB 9; Length 789;
 Best Local Similarity 55.2%; Pred. No. 1.2e-05;
 Matches 153; Conservative 0; Mismatches 118; Indels 6; Gaps 2;

QY 533 CCAAGGCTGTGTTATTCATTTGGCTAAGACTTTGGCTTGTATGGGCTAATCAACA 592
 DB 485 CAAAAGAGCTTTGATTCAGCTAGCTAAATAATTTGGCAATGTAAGGCAAAAGCGCA 544
 QY 593 TCAGAGTTAATCTTTAAACCAAGTTACATCAAGCTTGTGACCAAGATGTTATCA 652
 DB 545 TAAAGGCCAAGCGCTGTGGCTTAATGATCAATACCTCTGTCTCAATCTTATCTTG 604
 QY 653 ATGTAAAGCAAGATTTGACCAAGATGATCTTGTATCCCAACAAGAAATGTCG 712
 DB 605 AGG--ACGTCACTTCAAGAGGCAATGTTAGTAGACTCACTTGTGCTGTTGAG 661
 QY 713 AACCAAGGAATATCATTTGCTGTTTGTACTTGTCTTGAATGCTGCTTATACA 772
 DB 662 AGCCAAATGAAG--TTGATCATCACTAGTGGCTTCTTGTGCTTACTGCAAGCTTCTTATA 718
 QY 773 CTACTGTGCGCAGCTTACTGTTGATGTTGATGTTTAC 809
 DB 719 TTACTGTGTCAGACTATTGTTGTATGATGAGGTCCTCAC 755

RESULT 15

US-09-938-842A-948
 ; Sequence 948, Application US/09938842A
 ; Patent No. US20040009476A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 948
 ; LENGTH: 789
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana

US-09-938-842A-948

Query Match 6.9%; Score 56.2; DB 11; Length 789;
 Best Local Similarity 55.2%; Pred. No. 1.2e-05;
 Matches 153; Conservative 0; Mismatches 118; Indels 6; Gaps 2;

QY 533 CCAAGGCTGTGTTATTCATTTGGCTAAGACTTTGGCTTGTGATGGGCTAATCAACA 592
 DB 485 CAAAAGAGCTTTGATTCAGCTAGCTAAATAATTTGGCAATGTAAGGCAAAAGCGCA 544
 QY 593 TCAGAGTTAATCTTTAAACCAAGTTACATCAAGCTTGTGACCAAGATGTTATCA 652
 DB 545 TAAAGGCCAAGCGCTGTGGCTTAATGATCAATACCTCTGTCTCAATCTTATCTTG 604
 QY 653 ATGTAAAGCAAGATTTGACCAAGATGATCTTGTATCCCAACAAGAAATGTCG 712
 DB 605 AGG--ACGTCACTTCAAGAGGCAATGTTAGTAGACTCACTTGTGCTGTTGAG 661
 QY 713 AACCAAGGAATATCATTTGCTGTTTGTACTTGTCTTGAATGCTGCTTATACA 772
 DB 662 AGCCAAATGAAG--TTGATCATCACTAGTGGCTTCTTGTGCTTACTGCAAGCTTCTTATA 718
 QY 773 CTACTGTGCGCAGCTTACTGTTGATGTTGATGTTTAC 809
 DB 719 TTACTGTGTCAGACTATTGTTGTATGATGAGGTCCTCAC 755

Search completed: July 12, 2005, 15:54:44
 Job time : 640 secs

This Page Blank (usptg)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 11:47:19 ; Search time 3387 Seconds

(without alignment)
9170.488 Million cell updates/sec

Title: US-10-720-018-1

Perfect score: 816

Sequence: 1 atgactgactcattccaac.....atggtgttcactcttgg 816

Scoring table: IDENTITY NUC

Gapop 10.0 ; Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 276.6 | 33.9 | 899 | 6 | CNS077V0 |
| 2 | 129.2 | 15.8 | 969 | 9 | AL433186.T7 end of |
| 3 | 118.4 | 14.5 | 682 | 4 | CD457882 FG04d.02e |
| 4 | 113.4 | 13.9 | 713 | 7 | BI750180 FG02.10b0 |
| 5 | 113.4 | 13.9 | 836 | 7 | CF677471 CCA1590TF |
| 6 | 113.4 | 13.9 | 836 | 7 | CF701403 CCACT16TF |
| 7 | 112.8 | 13.8 | 777 | 7 | CF715948 CCA241TF |
| 8 | 112.8 | 13.8 | 787 | 7 | CF817896 CCA241TF |
| 9 | 111.2 | 13.6 | 738 | 7 | CF824542 EST701924 |
| 10 | 106.4 | 13.0 | 676 | 7 | CF812382 EST689764 |
| 11 | 102.8 | 12.6 | 718 | 1 | CF824543 EST701925 |
| 12 | 100.8 | 12.4 | 615 | 7 | AJ638396 |
| 13 | 97.2 | 11.9 | 447 | 7 | CO136295 |
| 14 | 93.2 | 11.4 | 511 | 1 | CO141255 EST835926 |
| 15 | 90.8 | 11.1 | 892 | 7 | AU249728 |
| 16 | 82.8 | 10.1 | 758 | 7 | CN811863 FG09.08p2 |
| 17 | 77.4 | 9.5 | 753 | 5 | CN812466 FG11.01e0 |
| 18 | 77.2 | 9.5 | 576 | 6 | BQ855471 QGB621.1 |
| 19 | 76.6 | 9.4 | 894 | 6 | CF820440 EST697822 |
| 20 | 76 | 9.3 | 833 | 7 | CD456321 FG03.03g0 |
| 21 | 76 | 9.3 | 839 | 7 | CO012226 EST800561 |
| 22 | 76 | 9.3 | 922 | 7 | CO012227 EST800562 |
| 23 | 76 | 9.3 | 949 | 7 | CO009933 EST798268 |
| 24 | 76 | 9.3 | 963 | 7 | CO004456 EST792791 |
| | | | | | CO004455 EST792790 |

| | | | | | |
|----|------|-----|------|---|----------|
| 25 | 75.8 | 9.3 | 692 | 5 | BQ992634 |
| 26 | 74.4 | 9.1 | 809 | 6 | CB905184 |
| 27 | 74.4 | 9.1 | 809 | 7 | CF876511 |
| 28 | 73.8 | 9.0 | 479 | 7 | CF191935 |
| 29 | 73 | 8.9 | 816 | 2 | BE642336 |
| 30 | 72.6 | 8.9 | 688 | 5 | BP098973 |
| 31 | 72.6 | 8.9 | 735 | 7 | CO140332 |
| 32 | 71.4 | 8.8 | 912 | 7 | CO009934 |
| 33 | 70.8 | 8.7 | 690 | 7 | CK569564 |
| 34 | 70.8 | 8.7 | 1048 | 9 | CNS07622 |
| 35 | 70.2 | 8.6 | 712 | 6 | CD458386 |
| 36 | 70 | 8.6 | 482 | 7 | CF190316 |
| 37 | 68 | 8.3 | 734 | 7 | CF447731 |
| 38 | 67.4 | 8.3 | 659 | 1 | AJ638576 |
| 39 | 67.4 | 8.3 | 971 | 9 | CNS07CVV |
| 40 | 66.8 | 8.2 | 800 | 7 | CO006699 |
| 41 | 65.4 | 8.0 | 786 | 5 | BQ855467 |
| 42 | 65.2 | 8.0 | 868 | 7 | CO006700 |
| 43 | 64.6 | 7.9 | 700 | 7 | CK567406 |
| 44 | 63.8 | 7.8 | 621 | 5 | BU014570 |
| 45 | 63.6 | 7.8 | 703 | 5 | BQ861127 |

ALIGNMENTS

| | | | | | |
|------------|---|-------------|-----|--------|-----------------|
| RESULT 1 | CNS077V0 | 899 bp | DNA | linear | GSS 08-JUL-2001 |
| LOCUS | CNS077V0 | | | | |
| DEFINITION | T7 end of clone BB0AA014G1 of library BB0AA from strain CBS 4732 | | | | |
| ACCESSION | AL433186 | | | | |
| VERSION | AL433186.1 | GI:12216600 | | | |
| KEYWORDS | GSS. | | | | |
| SOURCE | Pichia angusta | | | | |
| ORGANISM | Pichia angusta | | | | |
| REFERENCE | Bolotin-Fukuhara, M., Ben, E., Broctier, P., Casaregola, S., de-Montigny, J., Dujon, B., Durren, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuvéglise, C., Ozler-Kalageropoulos, O., Potier, S., Saurin, W., Tekala, F., Toffano-Nicoche, C., Wesolowski-Louvel, M., Wincker, P., and Weisenbach, J. | | | | |
| AUTHORS | Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies | | | | |
| TITLE | Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies | | | | |
| JOURNAL | FEBS Lett. 487 (1), 3-12 (2000) | | | | |
| PUBMED | 20584711 | | | | |
| REFERENCE | 2 (bases 1 to 899) | | | | |
| AUTHORS | Blandin, G., Llorente, B., Malpertuy, A., Wincker, P., Artiguenave, F. and Dujon, B. | | | | |
| TITLE | Genomic exploration of the hemiascomycetous yeasts: 13. Pichia angusta | | | | |
| JOURNAL | FEBS Lett. 487 (1), 76-81 (2000) | | | | |
| PUBMED | 20584723 | | | | |
| REFERENCE | 3 (bases 1 to 899) | | | | |
| AUTHORS | Genoscope. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (08-SBP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: sequef@genoscope.cns.fr Web: www.genoscope.cns.fr) | | | | |
| COMMENT | This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of | | | | |

```

FEATURES             the other extremity of this insert.
                        location/Qualifiers
                        1..899
                        /organism="Pichia angusta"
                        /mol_type="genomic DNA"
                        /strain="CBS 4732"
                        /db_xref="taxon:4905"
                        /clone="HB0A014G11"
                        /clone_lib="BROAD"
                        /note="end : 17"
misc_feature
<18..5633
/note="similar to P50167 [ D-arabinitol 2-dehydrogenase
ARDH ] [ Pichia stipitis ]
1 putative frameshift(s)"
/evidence=not experimental
complement(<752..5812)
/note="similar to Saccharomyces cerevisiae ORF YKL126w [
YPK1 ; ser/thr-specific protein kinase ]"
evidence=not experimental
complement(<733..5812)
/note="similar to Saccharomyces cerevisiae ORF YMR104c [
YPK2 ; ser/thr protein kinase ]"
evidence=not experimental

```

| Query Match | Similarity | 33.9% | Score 276.6 | DB 9 | Length 899 |
|-------------|------------|---|--------------|------|----------------|
| Best Local | Matches | 416 | Conservative | 0 | Mismatches 210 |
| | | | | | Indels 4 |
| | | | | | Gaps 1 |
| QY | 191 | TGAACTGAAAGAGCTCCAAAGATGGGTTTCATATGCTGTGATATTTCTGATCTGATA | 250 | | |
| DB | 4 | TCAAAGAGGATCTCGTGCACCAAGATGACACTGTAAGTTTGACATTAAGCCAGTGGAGC | 63 | | |
| QY | 251 | CCGTTCAAGAAGTGTGTTGGTCAAGTTCGTAAGATTTGGTGAAGTTGCCATGCACTTGG | 310 | | |
| DB | 64 | AGGTAAACGAGGCTCTTTGGTCTCAAAATATACGCCGACTTTGGAAGTACCCGCTGCAATCTG | 123 | | |
| QY | 311 | TTAACAACAGCTGTTACTGTGAAAACTTCCATGTGAATATACCAGCCAGCAAGAACGCTG | 370 | | |
| DB | 124 | TCAACACTGCTGGTTATTTGCGAAGAAATTTCCAGCCGAGCACTACCCGCGCAAGAACGCCG | 183 | | |
| QY | 371 | AGAAATGAGTGAAGTTAACTTTGGGTTCTTTGATATTTCTCAAGCTTTGCTTAAC | 430 | | |
| DB | 184 | AAAGAGCTTCTCAAGGTCACCTTTTGGGGCCCTGTGATCGTGCACAAAGATTCGCCAAAC | 243 | | |
| QY | 431 | CATTGATCAAGAAGGATCAAGGGGCTTCTGTTGTTTGAATGGTCTTATGTCTGATG | 490 | | |
| DB | 244 | CAGTTATCGATTAACAACATTCAGGAGAGGCTTCATCTGATGATCGGATTCATATGTCGGAG | 303 | | |
| QY | 491 | CCATTGTCAAGATCTCTAAAAACAAGTTGTCTACAACTGTCGAAGGCTGGTGTATTC | 550 | | |
| DB | 304 | AAATCGTCAACGCCCAACACGCCGATGGCTTACAAATGTCMAAGCCGGGTGATTC | 363 | | |
| QY | 551 | ATTGGCTTAAGACTTTGGCTGTGATGGGCTTAAGTACAACATCAGAGTTAATTTCTTAA | 610 | | |
| DB | 364 | ACATGCTCAAGTGTCTTCGCGCGCTGAATGGGCCAAATACAGATCAGATGTAATCTTGT | 423 | | |
| QY | 611 | ACCCAGTTACATCTACGGTCTTTGACCAAGAATGTATCAATGGTAAACGAAGATTGT | 670 | | |
| DB | 424 | CCGACAGATACATCTTACACCGCTCAACAAAGAACCTTATTAACGGTAAATACGAATGT | 483 | | |
| QY | 671 | ACAAACAGATGATCTCTGATATCCACAAAGAAATGCCGAACCAAGGAATACATGG | 730 | | |
| DB | 484 | ACCAAGAGATGGCTTTGCTGACACCCCAATGGGCCGCTTTTGGAGCTTAAGGAATTAACCG | 543 | | |
| QY | 731 | GGGCTTTTGTACTGCTTTCTGAATCTGCTCTTCATACACTACTGTGTCCAGCTTA | 789 | | |
| DB | 544 | GATACGATCTCTACTGCTCTCCAACTCCGCTTCTTCATTACACACGGGAGCAAAATCAT | 603 | | |
| QY | 790 | ---CTGGTGAATGGTGGTTCACTTCTTGG | 816 | | |
| DB | 604 | TATTTGTCTGACCGTGGTTCACCTTGTGG | 633 | | |

| | | | | |
|-----------------|--|------|--------|-----------------|
| RESULT 2 | | | | |
| CD457882 | | | | |
| LOCUS | | | | |
| DEFINITION | 969 bp | mRNA | linear | EST 14-JUN-2004 |
| ACCESSION | CPD457882 | | | |
| VERSION | Fg04d_02e10_A | | | |
| KEYWORDS | Fg04 AAFc ECORC_Fusarium_graminearum_mycelium_grown_on_wheat_heads Gibberella zeae cDNA clone Fg04d_02e10, mRNA sequence. | | | |
| SOURCE | CPD457882 | | | |
| ORGANISM | CPD457882.1 GI:31372622 | | | |
| REFERENCE | EST. | | | |
| AUTHORS | Gibberella zeae Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella. 1 (bases 1 to 969) Ouellet,T., Koul,A., Dan,H., Harris,L.J., Chapados,J., Couroux,P., De Moors,A., Hattori,U., Lacroix,C., Masotti,M., Robert,L.S., Singh,J.A., Spiroff,D. and Tinker,N.A. Fusarium graminearum mycelium grown on wheat heads under high humidity conditions unpublished (2003) | | | |
| JOURNAL COMMENT | Contact: Ouellet, Therese Eastern Cereal and Oilseed Research Centre Agriculture and Agri-Food Canada Neelby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA 0C6, CANADA Tel: (613) 759-1658 Fax: (613) 759-1701 Email: ouelletr@agr.gc.ca. | | | |
| FEATURES | Location/Qualifiers | | | |

| Query Match | Best Local Similarity | Matches | 289; Conservative | 15.8%; Score 129.2; DB 6; Length 969; 57.1%; Pred. No. 1.1e-25; 3; Mismatches 206; Indels 8; Gaps 3 |
|-------------|-----------------------|---|-------------------|--|
| QY | 305 | ACTTGCTTAACACACACTGGTTACTGTGAAATCTCCCAATGTGAAGATTACCGACGACAGA | 364 | 1..969 /organism="Gibberella zeae" /mol_type="mRNA" /strain="DAOM_180378" /db_xref="taxon:5518" /clone="Fg04d_02e10" /tissue_type="mycelium" /dev_stage="Asexual" /lab_host="E. coli" /clone_1ib="Fg04 AAFc ECORC_Fusarium_graminearum_mycelium grown on wheat heads" /note="Vector: Bluescript SK+; Site 1; EcorI; Site 2; XhoI; Fusarium grown on wheat (cv. Roblin) under high humidity. cDNA made using Stratagene kit." |
| DB | 254 | ACCTGATCACCTCAGCTGGCTTACAGGAGAACTTGAAAGCCTTAACTRACCCATCGACC | 313 | |
| QY | 365 | ACGCTGAGAAAGATGCTGAAGTTAACTTGTTGGGTTCTTTGATGTTTCAAGCTTTG | 424 | |
| DB | 314 | GTCTCCGTAAGCTTTGGGCTGTTACGTTACGTTACATATCTCTTGACATCATGCTG | 373 | |
| QY | 425 | CTAAAGCAATTGAT-CAAAGAAAGATTCAAAGGCTGCTTGTTGTTGATGTTGTTATG | 483 | |
| DB | 374 | CCAGGCACTTGATGCAGAAAGAAAGGCTCTCTGATGATC---GTATATATGTTGAGCATG | 429 | |
| QY | 484 | TCTGTGTCATTGTCAACGATCTCAAAACAAAGTTGTCTACACATGTCGAAGCTGGT | 543 | |
| DB | 430 | TCCGATTCATGTCAACGTTCTCTACGCTCAGGCTCCCTTAATAAAGCCGCAAGCGGT | 489 | |
| QY | 544 | GTTATCCATTGGGCTTAAGACTTTGGCTTGATGATGGGCTTAAGTACAAACATCAGAGTTAAT | 603 | |
| DB | 490 | GTGGCGCATCTGCGCTCTCTTGCCGCTGCAATGGGCTCAGGCAAAACATCGAGTCAAC | 549 | |
| QY | 604 | TCTTTAAACCAAGTTACATCTACGGTCTTTGACCAAGATGTTATCAATGTAAAGAA | 663 | |
| DB | 550 | TGCATCTCTCCGGTATCATGTTGACTGCATCATCAGAAATCTTTACGACAAACCG | 609 | |

QY 664 GAATGTACACAGATGATCTGTATCCACACAAAGATGTCACCAACCAAGCA 723
 DB 610 GATTCACAGGCGCAATGACCTTCCTTATCCCGGCGCAATGGACACACAGGAC 669
 QY 724 TACATGTGCTGCTTTTGTACTGTTCTGATCTGCTGCTTCACTACCTAGTGC 783
 DB 670 CTCATGGGGCCCGTGGACATCCCTTATCAGA---TGCTTCTTGATGTAGCTGTC 726
 QY 784 AGCTTACTGTTGATGTGTTTAC 809
 DB 727 CACATCCGATGATGCGGCTACAC 752

RESULT 3
 LOCUS B1750180 682 bp mRNA linear EST 14-JUN-2004
 DEFINITION Fg02_10b03_R Fg02 AAFPC ECORC Fusarium graminearum mycelium
 Gibberella zeae cDNA clone Fg02_10b03, mRNA sequence.
 ACCESSION B1750180
 VERSION B1750180.1 GI:15771982
 KEYWORDS EST.
 SOURCE Gibberella zeae
 ORGANISM Gibberella zeae

REFERENCE 1 (bases 1 to 682)
 AUTHORS Harris,L.J., Glasco,T., Rochelleau,H., Allard,S., Chapados,J.,
 Couroux,P., De Moore,A., Hattori,J.I., Ouellet,T., Robert,L.S.,
 Slough,J.A., Sprott,D. and Tinker,N.A.
 JOURNAL Unpublished Sequence Tags from Fusarium graminearum mycelium
 COMMENT (2001)
 TITLE Contact: Harris, Linda J.
 JOURNAL Eastern Cereal and Oilseed Research Centre
 COMMENT Agriculture and Agri-food Canada
 Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
 CANADA
 Tel: (613) 759-1314
 Fax: (613) 759-6566
 Email: harris1@agr.gc.ca.

FEATURES
 SOURCE location/Qualifiers
 1..682

/organism="Gibberella zeae"
 /mol_type="mRNA"
 /strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg02_10b03"
 /issue_type="mycelial tissue"
 /dev_stage="asexual"
 /lab_host="E. coli (Sure cells)"
 /clone_lib="Fg02 AAFPC ECORC Fusarium graminearum mycelium"
 /note="Vector: Bluescript SK+/XhoI-ECOR1; Site 1: EcoRI;
 Site 2: XhoI; Mycelial tissue was collected from v8 agar
 plates after a growth period of 6-7 days at 25 C with 14
 hrs (F/U/V) day light exposure. Mycelia was ground in
 liquid nitrogen prior to storage at -80 C until RNA
 extraction. Directional cloning with 5' end of cDNA cloned
 into EcoRI site of pBluescript and 3' end of cDNA cloned
 into XhoI site of pBluescript (Stratagene, La Jolla, CA)."

ORIGIN
 Query Match 14.5%; Score 118.4; DB 4; Length 682;
 Best Local Similarity 52.1%; Pred. No. 1.2e-22;
 Matches 342; Conservative 8; Mismatches 295; Indels 11; Gaps 4;

QY 14 TTCAACTTTAGATTGATGCGCACTTAACCATTTGACAGGTCCTGTGCTTTAG 73
 DB 20 TGCCACCTTTTACCTGACGCGTATGTGGCATGTGCACCGGTGTCAGTGGTGG 79
 QY 74 CTGAAGCTTTAATCAAGGCTTTGTGGCTTACGCTTGCATTTGCTGTGATTCG 133
 DB 80 GTCTTGTATTAGGTCAGAGATGATATTCTGTGATTCACCTTGCTTGTATATTTW 139

QY 134 ACCAAGAAAGACTCTGCAACCAAGCCGAATACCAATA---CGCTACTGAAGAT 190
 DB 140 ATTAAGAAAGACAGAGAAGCACTAGCTGATATTGAAGATTCAAAAAGAAACC 199
 QY 191 TGAAGTTAAAGAAATTCACCAAGATGGGTATCATAGCCGTGATATTCTGATTA 250
 DB 200 CTCGAGCCCGACGAATCCCAAGAGTCACTGCCATTATGCTGATATCTATCTGAAT 259
 QY 251 CCGTTCACAGAGTGTCTTCTCAAGTTCGAAGATTTGTGAAGTTGCACTTGACCTGG 310
 DB 260 CTGTGAGGCTTGTATTACCGAGGTTGTTAAGAGACAGAAAGATGACA---ACCTGG 316
 QY 311 TTACACACGCTGCTTACTGTGAAACTTCCTCAATGTAAGATTCCACCAAGAACGCTG 370
 DB 317 TCACCTCAGCTGGCTTACAGGAACTTCGAAGCCGTTAATACBSCATCGACCGCTTCC 376
 QY 371 AGAAGATGATGAAGTTACTGTTGGTCTTTGTATAGTTTCTCAACCTTGTAGC 430
 DB 377 GTAGCTTGGAGCTGTATACGTTGACGTTACATATTTCTTGGCAATCAGTCGACAGG 436
 QY 431 CATTGAT-CAAGAAGATATCAAGGCTCTGTTGTTTGAATGTTCTATGCTGAT 489
 DB 437 ACTTATATCAAGAAAGAGGCTCTGTGATATC---GTCAATGTTTATGATGTCGGA 492
 QY 490 GCCATTGCAAGATCTTCAAAACCAAGTGTCTTCAACATGTCACAGGCTGTGTATC 549
 DB 493 TCCATTGTCAAGTGTCTTCAAGCTTCAAGCTTCAAGTGTCTTCAAGTGTGTGCGC 552
 QY 550 CATTGGCTAGACTTTGGCTTGTGAATGGCTTAAATGATCAATCAAGTTATTTCTTA 609
 DB 553 CAACTGCTGCTTGTGTTGGCTGGAATGGCTCAAGCAACATCCGATTAACGCAATC 612
 QY 610 AACCAGGTTACATCTACGCTCTTTGACCAAGATGTTATCAATGTAAGCAAGA 665
 DB 613 TCTCGGTTACATGTTGACTGCACTCACTAGAAAGATCTTGAACGACACACCGGA 668

RESULT 4
 LOCUS CF677471/1 713 bp mRNA linear EST 16-AUG-2004
 DEFINITION CCAT990TF C.neoformans strain JEC21 Cryptococcus neoformans var.
 ACCESSION CF677471
 VERSION CF677471.1 GI:41531630
 KEYWORDS EST.
 SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
 var. neoformans)
 ORGANISM Cryptococcus neoformans var. neoformans
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
 Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
 Filobasidiella.

REFERENCE 1 (bases 1 to 713)
 AUTHORS Loftus,B.
 TITLE End sequencing of clones from a full length enriched, normalized
 JOURNAL JEC21 cDNA library
 COMMENT unpublished (2003)
 OTHER ESTS: CCAT990TR
 CONTACT: Brendan Loftus
 TIGR

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: crypt@tigr.org
 Seq primer: TF.

FEATURES
 SOURCE location/Qualifiers
 1..713
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCAT1990"
 /clone_lib="C.neoformans strain JEC21"

/note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 713;
Best Local Similarity 54.6%; Pred. No. 3.5e-21;
Matches 278; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

308 TGGTTACACAGCTGTTACTGTGTAAGAACTCCCATGTGAAGATTACCAAGCCAGAACG 367
649 TCGTCACTGCTGCGGATTTGTCGAAACTTTGTCGCTCAAGATACCCCATCATAGA 590
368 CTGAGAAAGATGTGAAGTTAACTTTGGTCTTTGATGTTTCTCAAGCCTTGTCTA 427
589 TCAGAAAGCTGTGACATCAACATTATGGTACTGTGATTTGGCACTTGAAGCTGCCA 530
428 AGCCATTGATCAAGAAAGATTCAGAGGTGCTTGTGTTGATTGGTTCTATGCTG 487
529 AGCTTATGCTGAA-----GGTGGTTCATTACCTTGCTGCATCTATGAGCG 482
488 GTGCATTGTCAAGATCTCTCAAAACCAAGTTGTCTACACATGTCCAAAGCTGTGTTA 547
481 GTAGCATTTGTCAAGCTTCTCAACCTCAACCCCTTACACTTTTCCAAAGCTGTGTC 422
548 TCCATTGTGGCTAAGCTTTGGCTGTGTGAATGGCTTAAGTACACATCAAGTTAATT 607
421 GACACATGGCTGATCCCTGCGCGCTGAATGGCTCTCAAGGGTATCCGTGTCAACGCTC 362
608 TAAACCCAGGTTACATCTACAGCTCTTTGACCAAGAAATGTTATCAATGATACGAAG 667
361 TTAGTCCCGGTTTACGCTCTCAACCACTTGACTTAAGGTCAATCTGCACCCCAACCC 302
668 TGTACACAGATGATCTCTGGTATCCCAACAAGATGTCCGAACCAAGAAATCA 727
301 TCCGGAAGATGGCTCAACCGTATCCCATGGGTGATGGCCGACCCCTTCTGATCTCA 242
728 TTGGTCTGTTTGTACTTCTTCTGAATCTGCTGCTTCACTACTGCTGCTCACT 787
241 AGGGTCCGTCATTACTTGTGCTTGA---CAGCTCCAAATGACACCACTGTGCTGAGA 185
788 TACTGTTGATGATGTTTCACTTCTTG 816
184 TCATGATTGACGCGGTTTACACTTGTCTG 156

RESULT 5
CF701403/c 836 bp mRNA linear EST 16-AUG-2004
LOCUS CCACCT16T C.neoformans strain JEC21 Cryptococcus neoformans var.
DEFINITION
ACCESSION CF701403
VERSION CF701403.1 GI:41555562
KEYWORDS EST.
SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)
ORGANISM Cryptococcus neoformans var. neoformans
Bukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella
1 (bases 1 to 836)
Loftus, B.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Other_ESTS: CCAC16T
Contact: Brendan Loftus
TIGR
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208

Email: crypt@tigr.org
Seq primer: TF.
Location/Qualifiers
1. 836

FEATURES

Source

/organism="Cryptococcus neoformans var. neoformans"
/mol_type="mRNA"
/strain="JEC21"
/db_xref="taxon:40410"
/clone="CCACT16"
/clone_1lb="C.neoformans strain JEC21"
/note="Vector: pCMVSPORT6; Site 1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 836;
Best Local Similarity 54.6%; Pred. No. 3.7e-21;
Matches 278; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

308 TGGTTACACAGCTGTTACTGTGTAAGAACTCCCATGTGAAGATTACCAAGCCAGAACG 367
665 TCGTCACTGCTGCGGATTTGTCGAAACTTTGTCGCTCAAGATACCCCATCATAGA 606
368 CTGAGAAAGATGTGAAGTTAACTTTGGTCTTTGATGTTTCTCAAGCCTTGTCTA 427
605 TCAGAAAGCTGTGACATCAACATTATGGTACTGTGATTTGGCACTTGAAGCTGCCA 546
428 AGCCATTGATCAAGAAAGATTCAGAGGTGCTTGTGTTGATTGGTTCTATGCTG 487
545 AGCTTATGCTGAA-----GGTGGTTCATTACCTTGCTGCATCTATGAGCG 498
488 GTGCATTGTCAAGATCTCTCAAAACCAAGTTGTCTACACATGTCCAAAGCTGTGTTA 547
497 GTAGCATTTGTCAAGCTTCTCAACCTCAACCCCTTCAACTTTTCCAAAGCTGTGTC 438
548 TCCATTGTGGCTAAGCTTTGGCTGTGTGAATGGCTTAAGTACACATCAAGTTAATT 607
437 GACACATGGCTGATCCCTGCGCGCTGAATGGCTCTCAAGGGTATCCGTGTCAACGCTC 378
608 TAAACCCAGGTTACATCTACAGCTCTTTGACCAAGAAATGTTATCAATGATACGAAG 667
377 TTAGTCCCGGTTTACGCTCTCAACCACTTGACTTAAGGTCAATCTGCACCCCAACCC 318
668 TGTACACAGATGATCTCTGGTATCCCAACAAGATGTCCGAACCAAGAAATCA 727
317 TCCGGAAGATGGCTCAACCGTATCCCATGGGTGATGGCCGACCCCTTCTGATCTCA 258
728 TTGGTCTGTTTGTACTTCTTCTGAATCTGCTGCTTCACTACTGCTGCTCACT 787
257 AGGGTCCGTCATTACTTGTGCTTGA---TACACCACTGTGCTGAGA 201
788 TACTGTTGATGATGTTTCACTTCTTG 816
200 TCATGATTGACGCGGTTTACACTTGTCTG 172

RESULT 6
CF715948/c 836 bp mRNA linear EST 16-AUG-2004
LOCUS CCAB241T C.neoformans strain JEC21 Cryptococcus neoformans var.
DEFINITION
ACCESSION CF715948
VERSION CF715948
KEYWORDS EST.
SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)
ORGANISM Cryptococcus neoformans var. neoformans
Bukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella
1 (bases 1 to 836)
Loftus, B.
REFERENCE
AUTHORS

TITLE End sequencing of clones from a full length enriched, normalized JEC21 cDNA library
JOURNAL Unpublished (2003)
COMMENT Other ESTs: CCAB241TR
 Contact: Brendan Loftus
 TIGR
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: crypt@tigr.org
FEATURES
 source
 location/Qualifiers
 1..836
 /organism="Cryptococcus neoformans var. neoformans"
 /mol_type="mRNA"
 /strain="JEC21"
 /db_xref="taxon:40410"
 /clone="CCAB241"
 /clone_1lb="C.neoformans strain JEC21"
 /note="Vector: pCMVSPORT6; Site 1: NotI EcoRV. The full length, normalized library was prepared from a variety of conditions using RNA provided by Joseph Heitman and Jennifer Lodge"

ORIGIN

Query Match 13.9%; Score 113.4; DB 7; Length 836;
 Best Local Similarity 54.6%; Pred. No. 3.7e-21;
 Matches 278; Conservative 0; Mismatches 216; Indels 15; Gaps 2;

308 TGGTTAACAAGCTGTTACTGTAAGAACTTCCCATGTAAGATTACCAAGCAAGACG 367
 670 TCGTCACTGCTCCCGGTATTCGAAACTTGTGCTCAGCATGCCATCGATTAGA 611
 368 CTGAGAAGATGGAAGTTAATTGTTGTTTGTATGTTTCTCAAGCTTGTCTA 427
 610 TCAAGAGCTGTGGACATCAATTAATGTTGTTGCGACTTGAAGCTGCCA 551
 428 AGCCATTGATCAAGAGATCAAGAGGTCTTCTGTTGTTTGAATGTTCTATGCTG 487
 550 AGCTTATGCTGAA-----GGTGTTCATTACCTCGCGCATCTATGAGCG 503
 488 GTGCCATTGTCAAGATCTCTCAAAACAAAGTTGTCTACCAATGTCGAAGCTGTGTTA 547
 502 GTAGATTGTCAACCTTCTCAACCTCAAAACCTTCAACCTTCAAGAGCTGTGTC 443
 548 TCCATTGGCTAAGCTTGGCTTGTGAATGGGCTAAGTCAACATCAAGTTAATCTT 607
 442 GACACATGGCTGATCCCTCGCGGTGAATGGGCTCTCAAGGATTCCTGTCAACGCTC 383
 608 TAAACCAAGTTACATCTACGCTCTTGAACCAAGATGTATCAATGTTAAGAAAT 667
 382 TTAGTCCGGTTACCTCTCAACCACTTACCTAAGTCTTCTCAAGCGCAACCCGTTTC 323
 668 TGTACAACAGATGATCTCTGTGATCCCAACAAAGATGTCCGAACCAAGAAATCA 727
 322 TCCGTAAGAGTGGCTCAACCGTATCCCATGGGTGAATGGCCGACCTTGTATCTCA 263
 728 TTGGGCTGTTTGAATCTTCTTGTGAATCTGCTGCTTCAATCACTATGTTGCAAGT 787
 262 AGGGGCGGTCACTTACCTGCTTGTGAAGCTCAAG---TACACCACTGTGTGTAGA 206
 788 TACTGTGATGATGTTTCACTTCTTGG 816
 205 TCATGATGTAAGCGCGTTACACTTCTTG 177

RESULT 7
 CF817896/c 777 bp mRNA linear EST 01-APR-2004
 LOCUS CF817896
 DEFINITION EST695278 Coccidioides posadasii saprobic phase cDNA library, greater than 4kb Coccidioides posadasii cDNA clone CIB147 5' end, mRNA sequence.

ACCESSION CF817896
VERSION CF817896.1 GI:45923774
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onygenales; mitosporic Onygenales; Coccidioides.
REFERENCE
 1 (bases 1 to 777)
 Gardner, M.J. and Cole, G.T.
 Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags
 Unpublished (2003)
JOURNAL Other ESTs: EST695277
COMMENT Contact: Gardner MJ
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301 838 3519
 Fax: 301 838 0208
 Email: gardner@tigr.org
 Seq primer: M13 Reverse.
FEATURES
 source
 location/Qualifiers
 1..777
 /organism="Coccidioides posadasii"
 /mol_type="mRNA"
 /strain="C735"
 /db_xref="taxon:199306"
 /clone="CIB147"
 /dev_stage="saprobic phase (mycelia)"
 /lab_host="E. coli DH10B, T1 phage resistant"
 /clone_1lb="Coccidioides posadasii saprobic phase cDNA library, greater than 4kb"
 /note="Vector: pXpress 1; Site 1: Not I; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA > 4 kb"

ORIGIN

Query Match 13.8%; Score 112.8; DB 7; Length 777;
 Best Local Similarity 54.1%; Pred. No. 5.3e-21;
 Matches 275; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

304 CACTGTGTTAACAAGCTGTTACTGTAAGAACTTCCCATGTAAGATTACCAAGCAAG 363
 708 CATCTGTCACATCTCGCGGCTTCACTGAAACCTTGAACCGCTGCTCAACCGCATAT 649
 364 AACGTGAAGATGTAAGTTAATTGTTGGTCTTGTATGTTTCTCAAGCTTT 423
 648 CGCATGCAAACTGTGCTGCTCAACGTTGATGATCTTTTGCAGTTCTGTT 589
 424 GCTAAGCATGATCAAGAGTATCAAGGCTCTCTGTTGTTTGAATGTTCTATG 483
 588 GCGAAACATGAT---GGCTGCAAGTGGCTGCAACATTTGTTTATGGAAGCATG 532
 484 TCTGTGCTATGTCACAGATCTCTCAAAACAAAGTTGTCTCAACATGTCCAGGCTGCT 543
 531 TCTGTGCTATGTTAAGTCTCTCAAGCAAGGCGCTCAATGTGCAAAAGGCTCG 472
 544 GTTATCAATTTGGCTAAGCTTGGCTGTGAATGGGCTAAGTCAACATCAAGATTAAT 603
 471 ATAAGGCACTTGGCTCTTCTGCACTTAATGGGCAAGTGGGAATCAAGATGAAC 412
 604 TCTTAAACCAAGTTACATCTACGCTCTTGAACCAAGATGTATCAATGTTAAGCA 663
 411 TGTATCACTCAAGATATATTTAGTCTGCTGACCGCAAGATTTCTTGACAGAACCTT 352
 664 GAATGTACAACAGATGATCTGTGATTCACCAACAAAGATGTCCGAACCAAGAA 723
 351 GATTCAAAGAGATGATCTCTCTCAATTCACAGCAAAATGGGTACACCTGAAGAC 292
 724 TACATGTGCTGTTTGTACTGCTTCTGAATCTGCTGTTCACTAATCACTAGTGTGCC 783
 291 CTCATGGCCCAAGTACCTTCTTATGATGA---TGGAGCAAAATATGTTACTGGGCA 235
 784 AGTTACTGTGATGTTGTTTCACTT 811

Db 234 GATTGAGGTTGATGGCGGTACCTT 207

RESULT 8
CF824542/c 787 bp mRNA linear EST 01-APR-2004
LOCUS EST701924 Coccidioides posadasii saprobic phase cDNA library, 2 to
DEFINITION 4 kb Coccidioides posadasii cDNA clone CIDAR90 3' end, mRNA
sequence.
ACCESSION CF824542.1 GI:45930599
VERSION EST.
KEYWORDS Coccidioides posadasii
SOURCE Coccidioides posadasii
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 787)
AUTHORS Gardner, M.J., and Cole, G.T.
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST701925
Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@ligr.org
Location/Qualifiers
1..787
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDAR90"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA
library, 2 to 4 kb"
/note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
Coccidioides posadasii saprobic phase cDNA library, size
fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 13.8%; Score 112.8; DB 7; Length 787;
Best Local Similarity 54.1%; Pred. No. 5.3e-21;
Matches 275; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

QY 304 CACTGGTTAACAAGCTGTTACTGTGAAACCTTCCCATGTAAGATTACCCAGC 363
DB 706 CATCTGGTACATCTGCGGGGTTCACTGAAACCTTTGACGCGGTGCCATCCGATGAT 647
QY 364 AACGTGAGAAGATGTTGAAGTTAATTGTTGGCTTTCTTGTATGTTTCTCAAGCTTT 423
DB 646 CCATGACGAAACCTGTCCTGCAAGTGAATGGTCAATCTTTTGCAGTTGCTGTT 587
QY 424 GCTAAGCCATGATCAAAAGAGTATCAAGGGTCTGTTGTTTGAATGGTCTATG 483
DB 586 GCAAAACATCTAT--GGCTGCAAGTCCCTGGCAGCATGTTGTTCAATGAAAGCATG 530
QY 484 TCTGGTGCATTGTCAACAGATCTCAAAACCAAGTTTCTTACAACATGTCGAAGCTGAT 543
DB 529 TCTGGTGCATTGTTAACGTGCTCAAGCAACAGCCCGTACATGCTGCAAGAGTGG 470
QY 544 GTTATTCATTGGCTTAAGACTTTGGCTTGTGAATGGGCTAAGTACAACTCAAGATTAT 603
DB 469 ATTAAGGCACTTGGCTCTCTTGGCAGTTGAATGGGCAAGTGGGAATCAGAGTGAAC 410
QY 604 TCTTTAAACCCAGTTACATCAAGTCTTTGACCAAGAAATGTTATCAATGTTAAGAA 663
DB 409 TGTATTCATGTCAGGATACATGTTGACTGCTGACCCGCAAGATCTTGAACAACCCCT 350

QY 664 GAATGTACAACAGATGATCTGTATCCCAACAAGAAATGTCGCAACGAAGAA 723
DB 349 GATCTCAAGAGAGATGACCTCCCTCATTCACAAGCAAAATGGGTATCACTGAAGAC 290
QY 724 TACATTTGCTGCTTTTGTACTTCTTGAATCTGCTGTTATCATCACTAGTGCC 783
DB 289 CTCATGGGCCCAAGACGTTCTTATTAAGTGA--TCCGACAAATATGTTACTGGGACA 233
QY 784 AGTTACTGTTGATGGTGGTTTCACTT 811
DB 232 GATTGAGGTTGATGGCGGTACACTT 205

RESULT 9
CF812382/c 738 bp mRNA linear EST 01-APR-2004
LOCUS EST689764 Coccidioides posadasii saprobic phase cDNA library,
DEFINITION greater than 4kb Coccidioides posadasii cDNA clone CIBA394 3' end,
mRNA sequence.
ACCESSION CF812382
VERSION EST.
KEYWORDS Coccidioides posadasii
SOURCE Coccidioides posadasii
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Onygenales; Mitosporic Onygenales; Coccidioides.
REFERENCE 1 (bases 1 to 738)
AUTHORS Gardner, M.J., and Cole, G.T.
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and
spherules via expressed sequence tags
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST689765
Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@ligr.org
Location/Qualifiers
1..738
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIBA394"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii saprobic phase cDNA
library, greater than 4kb"
/note="Vector: pExpress 1; Site_1: Not I; Site_2: Eco RV;
Coccidioides posadasii saprobic phase cDNA library, size
fractionated cDNA > 4 kb"

ORIGIN

Query Match 13.6%; Score 111.2; DB 7; Length 738;
Best Local Similarity 53.9%; Pred. No. 1.5e-20;
Matches 274; Conservative 0; Mismatches 228; Indels 6; Gaps 2;

QY 304 CACTGGTTAACAAGCTGTTACTGTGAAACCTTCCCATGTAAGATTACCCAGC 363
DB 612 CATCTGGTACATCTGCGGGGTTCACTGAAACCTTTGACGCGGTGCCATCCGATAT 553
QY 364 AACGTGAGAAGATGTTGAAGTTAATTGTTGGCTTTCTTGTATGTTTCTCAAGCTTT 423
DB 552 CCATGACGAAACCTGTCCTGCTCAACCTTATGATGCTATATCTTTTGCAGTTGCTGTT 493
QY 424 GCTAAGCCATGATCAAAAGAGTATCAAGGGTCTTGTGTTTGAATGGTCTATG 483
DB 492 GCAAAACATCTGAT--GGCTGCAAGTCCCTGGCAGCATGTTGTTCAATGGAAGCATG 436
QY 484 TCTGGTGCATTGTCAACAGATCTCAAAACCAAGTTTCTTACAACATGTCGAAGCTGAT 543
DB 435 TCTGGTGCATTGTTAAGTGTCTCAAGCAACAGCCGCTACATGCTGCAAGAGTTGG 376

QY 544 GTTATCCATTGGCTTAAGCTTTGGCTTGTGTAATGGGCTAAGTACAAATCAGAGTTAAT 603
DB 375 ATTAAGGCACTTGGCTTCTTCCCTTGGCAGTTGATGGGCAAGTGGGAATCAGAGTAAC 316
QY 604 TCTTAAACCCAGGTATCACTACAGTCCCTTTGACCAAGAAATGTTATCAATGTATACGA 663
DB 315 TGTATCAGTCCAGGAATCACTGTGACTGCTTACCCGCAAGATTCTTGACACCAACCT 256
QY 664 GAATGTACAAAGATGATCTCTGTATCCCAACAAAGATGTCCGAACCAAGGA 723
DB 255 GATCTCAAGAGAAATGACCTCCCTCANTCCAAAGGAATATGGTACACTTGAAAC 196
QY 724 TACATTGTCTGTTTGTACTTCTTCTGAATCTGCTCTTCAATACATCTGCTG 783
DB 195 CTCAATGGGCCCCAGTACAGTTCTTATGAGTGA---TGCGAACAATATGTTACTGGGGCA 139
QY 784 AGCTTACGTGTGATGTGCTTCACTT 811
DB 138 GATTGAGGTTGATGGCGGTACACTT 111

RESULT 10

LOCUS CF824543 676 bp mRNA linear EST 01-APR-2004
DEFINITION EST701925 Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb Coccidioides posadasii cDNA clone CIDAR90 5' end, mRNA sequence.

ACCESSION CF824543
VERSION CF824543.1 GI:45930600
KEYWORDS EST.
SOURCE Coccidioides posadasii
ORGANISM Coccidioides posadasii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

REFERENCE 1 (bases 1 to 676)
AUTHORS Gardner, M.J. and Cole, G.T.
TITLE Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags
JOURNAL Other ESTs: EST701924
COMMENT Contact: Gardner MJ
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@igr.org
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source

1..676
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDAR90"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_1lb="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"
/note="Vector: pEXpress 1; Site 1: Not I; Site 2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"

ORIGIN

Query Match 13.0%; Score 106.4; DB 7; Length 676;
Best Local Similarity 58.6%; Pred. No. 3.6e-19;
Matches 204; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 464 TTGTTTATGATGCTTCTATGCTGTGTCATGTCACAGATCTCAAAACCAAGTTGCT 523
DB 121 TTGCTTATGGAAGCAATGCTGTGCTGCTTGTAAAGTGCCTCAGCACAGGCCGT 180
QY 524 ACAACATGTCCAAAGGCTGTGTATCAATTGGCTAAGACTTTGGCTTGTGAATGGGCTA 583

DB 181 ACAATGCTGCAAAAGGCTCGATTAAGCACTTGGCTGCTTCTTGGCAGTTGAATGGCA 240
QY 584 AGTACACATCAGATTAATCTTAAACCCAGTTACATCAAGTCCCTTTGACCAAGA 643
DB 241 GTGTGGAAATCAAGATGAATCTATCACTCCAGATACATGTGACTGCTCCGACCCGA 300
QY 644 ATGTATCAATATGTAACAGAAATGTGAACAAGATGATCTGTATATCCCAACAA 703
DB 301 AGATTCTTGAAGCAACCTGATCTCAAGAGAAAGTGAACCTCCCTCATTTCCAAAGCA 360
QY 704 GAATCTCCGAACCAAGAAATATCATTTGTGCTGTTTGTACTTCTTCTGAATGTGCTG 763
DB 361 AAATGGTACACCTGAACACCTCATGAGCCCAAGTACCTTATATGATGA---TGCGA 417
QY 764 CTTCATACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811
DB 418 GCAATATGTTACTTGGGCAAGATTGAGGTTGATGGCGGTACACTT 465

RESULT 11

LOCUS AJ638396 718 bp mRNA linear EST 05-MAY-2004
DEFINITION AJ638396 Mgc Mycosphaerella graminicola cDNA clone mgc06e02f, mRNA sequence.

ACCESSION AJ638396
VERSION AJ638396.1 GI:47031453
KEYWORDS EST.
SOURCE Mycosphaerella graminicola
Mycosphaerella graminicola
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes et Chaetochytriumycetes incertae sedis; Mycosphaerellaceae; Mycosphaerella.

REFERENCE 1 (bases 1 to 718)
AUTHORS Keon, J.P.R., Hargreaves, J.A., Antoniw, J.F. and Hammond-Kosack, K.
TITLE Analysis of expressed sequence tags from the wheat fungal leaf blight pathogen, Mycosphaerella graminicola (anamorph Septoria tritici)
JOURNAL Fungal Genet. Biol. (2004) In press
COMMENT Contact: Keon J
Plant Pathogen Interactions Division,
Rothamsted Research,
Harpenden, Herts, UNITED KINGDOM
Tel: +44(0)1582 763133
Fax: +44(0)1582 760981
Email: john.keon@bbsrc.ac.uk
Insert Length: 800 Std Error: 100.00
Seq primer: M13 reverse.
Location/Qualifiers

FEATURES
source

1..718
/organism="Mycosphaerella graminicola"
/mol_type="mRNA"
/strain="Strit"
/db_xref="taxon:54734"
/clone="mgc06e02f"
/clone_1lb="Mgc"
/note="Vector: pSPORn1; library constructed from senescent wheat leaves 21-25 days after infection with Mycosphaerella graminicola exhibiting abundant hyphal growth and asexual sporulation"

ORIGIN

Query Match 12.6%; Score 102.8; DB 1; Length 718;
Best Local Similarity 53.0%; Pred. No. 4e-18;
Matches 268; Conservative 0; Mismatches 232; Indels 6; Gaps 2;
QY 304 CACTTGTAAACAGCTGTGTAAGAAATCTCCATGTGAATTAACCAAGCCAG 363
DB 49 CACCTGCTCATTCGCGTGGCTTACCGAGAACTTCAGCGCATGAGTACCAAGAC 108
QY 364 AACGCTGAAGATGTAAGATTAATCTTGGGTTCTTGTATGTTTCTCAAGCCTT 423
DB 109 CGCATGCAAGACTTTGGGAGTCAAGTCCACGCAATATCATCTGCTTGTGACCGTG 168

QY 424 GCTAAGCATTGATCAAGAGATATCAAGGAGCTTCTGTTGTTTATGTTTATG 483
| | | | |
Db 169 GCCAAGC---ATCTCATGAGAGGCAAGGCTTCTGCAACGATGATGATGATGATG 225
| | | | |
QY 484 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 543
| | | | |
Db 226 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 285
| | | | |
QY 544 GTTATTCATTTGGCTTGAAGCTTGTGATGATGATGATGATGATGATGATGATG 603
| | | | |
Db 286 GTCCGCAATCTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
| | | | |
QY 604 TCTTTAAACCAAGTTACATCTACGCTCTTTCAGCAAGAAATGTTATGATGATG 663
| | | | |
Db 346 TGCATCTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 405
| | | | |
QY 664 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723
| | | | |
Db 406 GAGCTGCGCAACAGTGAACGAGCTCTCATCCACAGGCAAGATGGTCCGCAAGAAC 465
| | | | |
QY 724 TACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 783
| | | | |
Db 466 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 522
| | | | |
QY 784 AGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 809
| | | | |
Db 523 GATCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 548
| | | | |

RESULT 12

LOCUS

CO136295 615 bp mRNA linear EST 17-JUN-2004

DEFINITION EST130966 Aspergillus flavus Normalized cDNA Expression Library
Aspergillus flavus cDNA clone NAFB016 5' and similar to
D-arabinicol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250)
(ARH) [yeast] [Candida tropicalis], mRNA sequence.ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMCO136295.1 GI:48885273
EST.
Aspergillus flavusAspergillus flavus
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT1 (bases 1 to 615)
Yu, J., Whitelaw, C.A., Nieman, W.C., Bhattacharya, D. and Cleveland, T.E.
Aspergillus flavus expressed sequence tags for identification of
genes with putative roles in aflatoxin contamination of crops
FEMS Microbiol. Lett. (2004) in pressContact: Yu J
Food and Feed Safety Research Unit
USDA/ARS, Southern Regional Research Center
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
Tel: 504 286 4405
Fax: 504 286 4419
Email: jinyu@arrrc.ars.usda.gov
Contact Dr. Yu at USDA/ARS SRRC (jinyu@arrrc.ars.usda.gov) for clone
informationPCR Primers
FORWARD: M13F
BACKWARD: M13R
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

Location/Qualifiers
1..615

/organism="Aspergillus flavus"

/mol_type="mRNA"

/strain="NRRL 3357"

/db_xref="taxon:5059"

/clone="NAFB016"

/sex="asexual mycelia"

/cell_type="mycelia"

/dev_stage="developmental stages from 18 to 96 hours"

/lab_host="E. coli DH10B T1 resistant cells"
/clone_id="Aspergillus flavus Normalized cDNA Expression
Library"
/note="Vector: pBluescript (SK+) (Stratagene), antibiotic
selection marker: Carbenicillin; Site 1: NotI, at the 5
prime end; Site 2: EcoRI, at the 3 prime end; This
normalized cDNA expression library was constructed using a
mixture of mycelial cells grown under eight different
medium conditions and harvested at 5 time points (18, 24,
48, 72, 96 hours). The poly-A sequence was trimmed off
before ligating to vector."

ORIGIN

Query Match 12.4%; Score 100.8; DB 7; Length 615;
Best Local Similarity 51.2%; Pred. No. 1.4e-17;
Matches 287; Conservative 0; Mismatches 267; Indels 6; Gaps 2;

QY 196 TTGAAGAAGTTCGAAGATGGGTTCATATGCTGATGATGATGATGATGATGATG 255
| | | | |
Db 38 TTGAAGAAGTTCGAAGATGGGTTCATATGCTGATGATGATGATGATGATGATG 97
| | | | |
QY 256 CACAAGGTGTTGCTCAAGTGTCAAGATTTTGTGATGATGATGATGATGATGATG 315
| | | | |
Db 98 AACGATGCTCTCTCGATATATATCTCAAGCAGGCAAGAT---GCAACATGATGATG 154
| | | | |
QY 316 AAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375
| | | | |
Db 155 TCCGCGGATTCAGCGAAGATCTGATGATGATGATGATGATGATGATGATGATG 214
| | | | |
QY 376 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
| | | | |
Db 215 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 271
| | | | |
QY 436 ATCAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495
| | | | |
Db 272 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 331
| | | | |
QY 496 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 555
| | | | |
Db 332 GTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 391
| | | | |
QY 556 GCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615
| | | | |
Db 392 GCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 451
| | | | |
QY 616 GATTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 675
| | | | |
Db 452 GATTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511
| | | | |
QY 676 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 735
| | | | |
Db 512 AAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 571
| | | | |
QY 736 GTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 795
| | | | |
Db 572 GTTACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
| | | | |

RESULT 13

LOCUS

CO141255 447 bp mRNA linear EST 17-JUN-2004

DEFINITION EST130926 Aspergillus flavus Normalized cDNA Expression Library

Aspergillus flavus cDNA clone NAFB016 5' end, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMCO141255.1 GI:48890246
EST.
Aspergillus flavusAspergillus flavus
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.REFERENCE
AUTHORS
TITLE
Yu, J., Whitelaw, C.A., Nieman, W.C., Bhattacharya, D. and Cleveland, T.E.
Aspergillus flavus expressed sequence tags for identification of

JOURNAL COMMENT

Genes with putative roles in aflatoxin contamination of crops
FEMS Microbiol. Lett. (2004) In press
Contact: Yu J
USDA/ARS, Southern Regional Research Center
1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA
Tel: 504 286 4405
Fax: 504 286 4419
Email: jiyu@arrrc.ars.usda.gov
Contact Dr. Yu at USDA/ARS SRRC (jiyu@arrrc.ars.usda.gov) for clone
information
PCR Primers
FORWARD: M13F
BACKWARD: M13R
Seq primer: M13 Forward
POLYA=No.

FEATURES

source Location/Qualifiers

1..447
/organism="Aspergillus flavus"
/mol_type="mRNA"
/strain="NRRL 3357"
/db_xref="taxon:5059"
/clone="NAPEA37"
/sex="asexual mycelia"
/cell_type="mycelia"
/dev_stage="developmental stages from 18 to 96 hours"
/lab_host="E. coli DH10B T1 resistant cells"
/clone_1ib="Aspergillus flavus Normalized cDNA Expression library"
/note="vector: pBluescript (SK+) (stratagene), antibiotic selection marker: Carbenicillin, Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

ORIGIN

Query Match 11.9%; Score 97.2; DB 7; Length 447;
Best Local Similarity 58.7%; Pred. No. 1.4e-16;
Matches 168; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 464 TTGTTTGAATGTTCTATGTCGTGTCATGTCACGATCTCAAAACCAAGTTGCT 523
Db 162 TTGTCATGATGTTGTCATGTCGTGTCATGTCACGTCGCGACCGCTCCTT 221
Qy 524 ACAACATGTCAGGCTGTGTTATCCATTGGCTAAGACTTGGCTTGATGGGCTA 583
Db 222 ACAACGCGCCAGCGCTGTTCCTCAACTGCGCGCTCTTCCGTCGATGGGCG 281
Qy 584 AGTACAATCATGAGTTAATCTTTAAACCAAGTTACGATCAAGTCTTTGACAGA 643
Db 282 GTCAAGACATCCGGTGAATGTCATCAAGCCGGAATACGTTACTGCCCTCCGCA 341
Qy 644 ATGTTATCATGTTAACAAGATTTGTAACAAGATGATCTGTATCCACAACAA 703
Db 342 AGATTTTGATGAGAACCCCGAATGGGGACAATGATGATCTGCTATCCCAACGCA 401
Qy 704 GAATGTCGAAACCAAGATATCATGTTGCTGTTTGTACTTGGT 749
Db 402 AGATGGTACTCCGAGGACCTGATGGGTCCGTTACTCTCTGCT 447

RESULT 14

AU249728 511 bp mRNA linear EST 22-APR-2004
LOCUS AU249728 SL Lolium multiflorum cDNA clone SL002A09-5, mRNA
DEFINITION Sequence.
ACCESSION AU249728
VERSION AU249728.1 GI:46506997
KEYWORDS EST.
SOURCE Lolium multiflorum (Italian ryegrass)

ORGANISM

Lolium multiflorum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Poace; Lolium.

1 (bases 1 to 511)

Ikedu,S.

Lolium multiflorum EST Project

Unpublished (2004)

Contact: Seishi Ikeda

Japan Grassland Farming Forage Seed Association (JFSA)

Forage Crop Research Institute (FCRI)

Higashikada 388-5, Nishinasuno, Tochigi 329-2742, Japan

Tel: 81-287-37-6755

Fax: 81-287-37-6757

Email: sikedu67@fjfas.or.jp

Contact:Tadashi Takamizo (takamizo@affrc.go.jp)

National Institute of Livestock and Grassland Science, Nishinasuno

Resistance gene analog.

Location/Qualifiers

FEATURES

source

1..511
/organism="Lolium multiflorum"
/mol_type="mRNA"
/db_xref="taxon:4521"
/clone="SL002A09-5"
/tissue_type="Powdery mildew infected leaf"
/clone_1ib="SL"

ORIGIN

Query Match 11.4%; Score 93.2; DB 1; Length 511;
Best Local Similarity 58.5%; Pred. No. 2.1e-15;
Matches 200; Conservative 0; Mismatches 138; Indels 4; Gaps 2;

Qy 469 TTGATGTTCTATGTTCTGTCGTCATGTCACGATCTCAAAACCAAGTTGTCACAC 528
Db 2 TTTATGTAACATGTCAGGACCGTCGTGAATGTTCCACAACCAAGCACCATTATAC 61
Qy 529 ATGTCCAGGCTGTGTTATTCATTGTCGTTAAGACTTGGCTTGATGGGCTAAGTAC 588
Db 62 GCCGTAAAGCTGCAATTCGCGACCTGGCATCGATGATGAGATGGAGGCTCATGCA 121
Qy 589 AACATCAAGTTAATCTTTAAACCAAGTTAATCTGATCGCTCTTTGACCAAGATTT 648
Db 122 GAATTCAGTCAATGATGATTCCTGATGATGATGATGATGATGATGATGATGAT 181
Qy 649 A-TCAATGTCACGAAGATTTGTAACAAGATGATCTGATGATGATGATGATGAT 707
Db 182 ACTGAAAGATCTCTGATTTAAAGAACATGATGATGATGATGATGATGATGATGAT 241
Qy 708 GTCCGAACCAAGAAATACATGTCGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 767
Db 242 GGGTATTCAGAAATTTGATGAGGAGGAGGATTTCTTGTGCTTCTGA---TGCTATC 298
Qy 768 ATACATGACGGCGCTGATCTAAGATGATGATGATGATGATGATGATGATGATGAT 809
Db 299 ATACATGACGGCGCTGATCTAAGATGATGATGATGATGATGATGATGATGATGAT 340

RESULT 15

CN811863/c 892 bp mRNA linear EST 01-JUN-2004
LOCUS CN811863 Fg09_08p20 A Fg09 AAF C EORC Fusarium graminearum simple_substrate
DEFINITION Gibberella zeae cDNA clone Fg09_08p20, mRNA sequence.
ACCESSION CN811863
VERSION CN811863.1 GI:47835874
KEYWORDS EST.
SOURCE Gibberella zeae
ORGANISM Gibberella zeae
DEFINITION Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocerales; Nectriaceae; Gibberella.
REFERENCE 1 (bases 1 to 892)
Watson,R.J., Heyes,R., Couroux,P., De Moers,A., Harris,L.J.,
Hattori,J., Lacroix,C., Masotti,M., Ouellet,T., Robert,L.S.,
Singh,J.A., Spott,D. and Tinker,N.A.

TITLE A cDNA library prepared from *Fusarium graminearum* grown on a simple

substrate

JOURNAL Unpublished (2003)

COMMENT Contact: Watson, Robert.J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA

Tel: (613) 759-1655

Fax: (613) 759-1701

Email: watsonrj@agr.gc.ca.

FEATURES Location/Qualifiers

source

1..892

/organism="Gibberella zeae"

/mol_type="mRNA"

/strain="DAOM 180378"

/db_xref="taxon:5118"

/clone="Fg09_08p20"

/tissue_type="Mycelium"

/dev_stage="Asexual"

/lab_host="E. coli DH10B"

/clone_11b="Fg09_AAFc_ECORC_Fusarium graminearum_simple_substrata"

/note="Vector: pBluescript II+, Site 1: EcoRI; Site 2:

XhoI; Fusarium graminearum grown on a simple substrate--

minimal media supplemented with amino acids."

ORIGIN

Query Match 11.1%; Score 90.8; DB 7; Length 892;

Best Local Similarity 54.5%; Pred. No. 1.2e-14;

Matches 204; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

```

QY 437 TCAAGAAGGATGTCAGAGGCTCTTCTGTTGTTTGAATGTTCTATGCTGGTGCCATTG 496
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 TCATGAGAGGTGAGTGCCCGGTAGTATGTTGGTATGGAAGCAATGCTGGTGCCATTG 329
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 497 TCAACGATCCTCAAAACCAAGTTGTCTACAACATGTCACAGGCTGGTATTCATTGG 556
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 TCATATGCCACAGCAAGCCCATATTAAGCGCGCAAAAGCACTGTTGACACCTGG 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 557 CTAGACTTTGGCTGTGTAATGGCTTAAGTACACATCAGAGTTAATCTTTAAACCAAG 616
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 CTGCTTCTCTTCAAGTGAGTGAGGCTCAAGCTGAATCCGTGTCAATGATCTCTCTG 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 617 GTTACATCTACAGTCTTTGACCAAGATGTTATCATGTGTAAAGAAATTGTACAACA 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 GCTATATGTTGACTGTTTGACACAGAAAGATCCTGACGACAAACCTGATTTAGAGAGAA 149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 677 GATGATCTCTGGTATCCACAAACCAAGATGTCGAACCAAGGATATCATTTGGTGTG 736
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 CCTGACATCCTCATTTCTCAAGGTCGATGGAGTCTCAAGATTTGATGGAGCCGG 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 737 TTTTGAATCTTTCTGTAATCTGCTGTTCAATACATACTAGTGGCAGCTTACTGGTTG 796
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 TAACCTTTCTGCTATCAGA--TGCGTCACTTATATGACTGGGGCAGATGTTCCAGTTG 32
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 797 ATGGTGGTTTCACT 810
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31 ATGGAGATATCACT 18
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: July 12, 2005, 14:20:51
Job time : 3392 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 9, 2005, 10:15:36 ; Search time 162 Seconds

(without alignments)
649.375 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415
Sequence: 1 MTDYIPFRFDGHLTVTGA.....SAASYTGASLLVDGFTSW 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 366760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980a:*
2: geneseqp1980a:*
3: geneseqp2000a:*
4: geneseqp2000a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003a:*
8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1415 | 100.0 | 272 | 8 | ADQ94602 Ambrosioz |
| 2 | 424.5 | 30.0 | 258 | 3 | AAV56815 Bactillu |
| 3 | 402.5 | 28.4 | 257 | 8 | ADN20270 Bactillu |
| 4 | 396 | 28.0 | 211 | 8 | ADS24003 Bactillu |
| 5 | 393 | 27.8 | 266 | 5 | AAE28702 Bactillu |
| 6 | 381 | 26.9 | 283 | 8 | ADJ98100 Ketoreduc |
| 7 | 380 | 26.9 | 283 | 8 | ADJ98118 Ketoreduc |
| 8 | 378 | 26.7 | 283 | 4 | AAW64777 C. magno |
| 9 | 378 | 26.7 | 283 | 2 | AAW64777 C. magno |
| 10 | 378 | 26.7 | 283 | 4 | AAW64777 C. magno |
| 11 | 378 | 26.7 | 283 | 8 | ADJ98066 Ketoreduc |
| 12 | 378 | 26.7 | 283 | 8 | ADJ98060 Ketoreduc |
| 13 | 378 | 26.7 | 283 | 8 | ADJ98064 Ketoreduc |
| 14 | 376 | 26.6 | 253 | 8 | ADS22515 Bactillu |
| 15 | 376 | 26.6 | 283 | 8 | ADJ98104 Ketoreduc |
| 16 | 376 | 26.6 | 283 | 8 | ADJ98062 Ketoreduc |
| 17 | 376 | 26.6 | 283 | 8 | ADJ98112 Ketoreduc |
| 18 | 374 | 26.4 | 245 | 8 | ADN20651 Bactillu |
| 19 | 374 | 26.4 | 283 | 8 | ADJ98102 Ketoreduc |
| 20 | 373 | 26.4 | 283 | 8 | ADJ98108 Ketoreduc |
| 21 | 373 | 26.4 | 283 | 8 | ADJ98114 Ketoreduc |
| 22 | 373 | 26.4 | 283 | 8 | ADJ98116 Ketoreduc |
| 23 | 370 | 26.1 | 283 | 8 | ADJ98110 Ketoreduc |
| 24 | 367 | 25.9 | 251 | 8 | ADS25925 Bactillu |
| 25 | 367 | 25.9 | 283 | 8 | ADJ98106 Ketoreduc |

| | | | | | |
|----|-------|------|-----|---|--------------------|
| 26 | 364 | 25.7 | 263 | 6 | ABJ26225 Aspergill |
| 27 | 362 | 25.6 | 251 | 8 | ADS25181 Bactillu |
| 28 | 358.5 | 25.3 | 266 | 7 | ADG25179 C. fulvum |
| 29 | 358.5 | 25.3 | 266 | 7 | ADG25185 C. fulvum |
| 30 | 358.5 | 25.3 | 281 | 8 | ADS42909 Bactillu |
| 31 | 358.5 | 25.3 | 319 | 8 | ADS24222 Bactillu |
| 32 | 357 | 25.2 | 245 | 8 | ADS25559 Bactillu |
| 33 | 354.5 | 25.1 | 266 | 7 | ADG25187 A. altern |
| 34 | 350 | 24.7 | 293 | 8 | ADN21234 Bactillu |
| 35 | 344 | 24.3 | 316 | 7 | ABO66473 Klebsiell |
| 36 | 343.5 | 24.3 | 294 | 6 | ABJ25625 Aspergill |
| 37 | 343 | 24.2 | 268 | 8 | ADS23247 Bactillu |
| 38 | 340.5 | 24.1 | 257 | 5 | ABP38717 Staphyloc |
| 39 | 340.5 | 24.1 | 257 | 8 | ADS05249 Staphyloc |
| 40 | 337.5 | 23.9 | 267 | 7 | ADG25186 C. herbar |
| 41 | 337.5 | 23.9 | 267 | 7 | ADG25184 C. herbar |
| 42 | 336 | 23.7 | 255 | 8 | ADN20291 Bactillu |
| 43 | 326.5 | 23.1 | 284 | 8 | ADN21004 Bactillu |
| 44 | 324.5 | 22.9 | 245 | 8 | ADN25229 Bactillu |
| 45 | 322.5 | 22.8 | 256 | 7 | ABO81728 Pseudomon |

ALIGNMENTS

| | | |
|----------|--|---------------------------------------|
| RESULT 1 | ADQ94602 | standard; protein; 272 AA. |
| ID | ADQ94602 | standard; protein; 272 AA. |
| XX | ADQ94602; | |
| DT | 23-SEP-2004 | (first entry) |
| DE | Ambrosiozyma monospora | NADH dependent L-xylulose reductase. |
| KW | NADH dependent L-xylulose reductase; fermentation; biomass; enzyme; | |
| KM | EC 1.1.1.10. | |
| OS | Ambrosiozyma monospora. | |
| XX | US2004132074-A1. | |
| PD | 08-JUL-2004. | |
| PF | 24-NOV-2003; 2003US-00720018. | |
| PR | 16-FEB-2001; 2001FI-00000308. | |
| PR | 15-FEB-2002; 2002WO-FI000125. | |
| PR | 10-MAR-2003; 2003US-00257821. | |
| PR | 12-SEP-2003; 2003FI-00001307. | |
| PA | (VALM) VALTTON TEKNIILINEN TUTKIMUSKESKUS. | |
| XX | Verho R, Richard P, Penttila M; | |
| DR | WPI; 2004-517001/49. | |
| XX | N-PSDB; ADQ94601. | |
| PT | New DNA molecule encoding NADH dependent L-xylulose reductase, useful in | |
| PT | producing fermentation products or in conversion of cheap biomass to | |
| PT | useful products. | |
| PS | Claim 26; SEQ ID NO 2; 15pp; English. | |
| CC | The present invention relates to a NADH dependent L-xylulose reductase | |
| CC | enzyme protein and its encoding polynucleotide. The invention is useful | |
| CC | in producing fermentation products or in conversion of cheap biomass to | |
| CC | useful products. The present sequence is an Ambrosiozyma monospora NADH | |
| CC | -1484 NADH dependent L-xylulose reductase, EC 1.1.1.10. | |
| XX | Sequence 272 AA; | |
| SO | Query Match | 100.0%; Score 1415; DB 8; Length 272; |

Best Local Similarity 100.0%; Pred. No. 1,1e-137;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDTYIPFRPDGHLTIYTGACGGLAEALIKGLIAYGSDIALDIDQEXTAKQAEYHKYA 60
DB 1 MTDTYIPFRPDGHLTIYTGACGGLAEALIKGLIAYGSDIALDIDQEXTAKQAEYHKYA 60

QY 61 TEEKLVKVPWKGSYACDISDPTVHKVFAQVADFGKLPPLHNTAGYCENPCEDYPA 120
DB 61 TEEKLVKVPWKGSYACDISDPTVHKVFAQVADFGKLPPLHNTAGYCENPCEDYPA 120

QY 121 KNAEKMYVNLGSLIYVSOAFAPKLIKGIKQASVVLIGSMGAIIVDPQNVYYNMSKA 180
DB 121 KNAEKMYVNLGSLIYVSOAFAPKLIKGIKQASVVLIGSMGAIIVDPQNVYYNMSKA 180

QY 181 GVHIAKTLACEMAKYKINRVNSLNPGYIYGPITKRVINGNEELYNRWISGIPQRMSEPK 240
DB 181 GVHIAKTLACEMAKYKINRVNSLNPGYIYGPITKRVINGNEELYNRWISGIPQRMSEPK 240

QY 241 EYIGAVLYLSESAASYTTGASLLVDGGFTSM 272
DB 241 EYIGAVLYLSESAASYTTGASLLVDGGFTSM 272

RESULT 2

AA56815
ID AA56815 standard; protein; 258 AA.
AC AA56815;
XX
XX
XX 31-MAR-2000 (first entry)
DT
XX Bacillus D-arabinitol dehydrogenase.
KW
XX D-arabinitol dehydrogenase; clinical diagnosis; mycosis.
XX
XX Bacillus sp.
OS
XX JP11332569-A.
PN
XX 07-DEC-1999.
PD
XX 26-MAY-1998; 98JP-00143637.
PF
XX 26-MAY-1998; 98JP-00143637.
XX
XX 26-MAY-1998; 98JP-00143637.
PR
XX (IKED-) IKEDA SHOKEN KK.
XX (NIPK) NIPPON KAYAKU KK.
XX
XX WPI: 2000-091353/08.
DR N-PSDB; AA246762, AA246763.
XX
XX Arabinitol dehydrogenase gene encoding D-arabinitol dehydrogenase -
PT useful as a clinical diagnosing agent for mycosis.
XX
XX Claim 1; Page 10-11; 14pp; Japanese.
XX
XX The invention relates to gene encoding D-arabinitol dehydrogenase,
CC isolated from Bacillus sp. IKD-5A868 strain. The protein can be expressed
CC by standard recombinant methodologies. D-arabinitol dehydrogenase is used
CC as a clinical diagnosing agent for mycosis. The present sequence
CC represents the D-arabinitol dehydrogenase enzyme
XX
XX Sequence 258 AA;

Query Match 30.0%; Score 424.5; DB 3; Length 258;
Best Local Similarity 38.0%; Pred. No. 5,1e-35;
Matches 103; Conservative 44; Mismatches 99; Indels 25; Gaps 7;

QY 8 FPFDEHLLITVTCAGCGLABALIKGLIAYGSDIALDI--DQEXTAKQAEYHKYATEELK 65
DB 7 FPLDGRVTVITVTCAGCGLABALIKGLIAYGSDIALDI--DQEXTAKQAEYHKYATEELK 65

QY 66 LKEVPMKGSYA---CDISDPTVHKVFAQVADFGKLPPLHNTAGYCENPCEDYPAK 121
DB 53 -NEIRSCGVEATVQVDTVDTEAQVEALIRQVEQYGRLDV-LVNNAGIVKDRABEELPYE 110

QY 122 MAEKMYVNLGSLIYVSOAFAPKLIKGIKQASVVLIGSMGAIIVDPQNVYYNMSKAG 181
DB 111 TWRNVNWNVNSVFLMSKAAGKRVITRG-KG-STINISSMSGLIWNTPQQAAYNVSKAG 168

QY 182 VIHIAKTLACEMAKYKINRVNSLNPGYIYGPITKRVINGNEELYNRWISGIPQRMSEPK 241
DB 169 VIHIAKTLACEMAKYKINRVNSLNPGYIYGPITKRVINGNEELYNRWISGIPQRMSEPK 228

QY 242 YIGAVLYLSESAASYTTGASLLVDGGFTSM 272
DB 229 LGIYAVLYASD-ASSFATGVFTIDGGYTTM 258

RESULT 3

ADN20270
ID ADN20270 standard; protein; 257 AA.
AC
XX
XX ADN20270;
XX
XX 02-DEC-2004 (first entry)
DT
XX
XX Bacterial polypeptide #2923.
DE
XX

Recombinant DNA construct; transformed plant; improved plant property;
cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
pathogen tolerance; pest tolerance; plant disease resistance;
cell cycle pathway modification; plant growth regulator;
homologous recombination; seed oil yield; protein yield; carbohydrate;
nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
bacterial polypeptide.

OS Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY) CAO Y.
XX (HINK) HINKLE G J.
XX (SLAT) SLATER S C.
XX (CHEN) CHEN X.
XX (GOLD) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI: 2004-061375/06.
XX

New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 2923; 122pp; English.

The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,

| | | |
|----------|---|--|
| KM | | ethanol; enzyme. |
| OS | | Hypocrea jecorina. |
| XX | | |
| PN | | WO200266616-A2. |
| XX | | |
| PD | | 29-AUG-2002. |
| PF | | 15-FEB-2002; 2002WO-FI000125. |
| XX | | |
| PR | | 16-FEB-2001; 2001FI-00000308. |
| XX | | |
| PA | (VALM) | VALTION TEKUNILINEN TUTKIMUSKESKUS. |
| PJ | Londesborough J, | Penttilae M, Richard P; |
| DR | WPI; 2002-691618/74. | |
| XX | N-PsDB; AAD46155. | |
| PT | | Genetically modified fungus for producing useful products such as |
| XX | | ethanol, lactic acid and xylitol, from biomass containing L-arabinose, |
| XX | | has increased ability to utilize L-arabinose. |
| PS | | Disclosure; Fig 3; 32pp; English. |
| XX | | |
| CC | The invention relates to genetically modified fungus with an increased | |
| CC | ability to utilize L-arabinose, where the fungus has been transformed | |
| CC | with a DNA sequence encoding an L-arabinitol 4-dehydrogenase (EC 1.1. | |
| CC | 1.12) or L-xylulose reductase (EC 1.1.1.10) or both the DNA sequences. | |
| CC | Genetically modified fungus is useful for producing useful products from | |
| CC | biomass containing L-arabinose. The useful product include ethanol, | |
| CC | lactic acid or xylitol preferably ethanol. It is also useful to ferment a | |
| CC | carbon source such as biomass comprising agricultural or forestry | |
| CC | products and waste products containing L-arabinose and also other | |
| CC | pentoses or other fermentable sugars. The present sequence is T. reesei L | |
| CC | -xylulose reductase. (Updated on 29-AUG-2003 to standardise OS field) | |
| XX | | |
| SQ | Sequence 266 AA; | |
| | | |
| | Query Match | 27.8%; Score 393; DB 5; Length 266; |
| | Best Local Similarity | 36.4%; Pred. No. 9, 9e-32; |
| | Matches 102; Conservative | 51; Mismatches 103; Indels 24; Gaps 8 |
| Oy | 1 MTDTYFT-----FRFDGHLTVTGACG--GLAEALIKGLAYSGDIALLDIDDEKTNA | 51 |
| Dd | : : : : : : : : : | |
| | 1 MEQPVTANRLIDLFLSKGVVVVVTASGPFGWGIEAARGCAEMGADLAITYTSRRKEAF | 60 |
| Oy | 52 KQAETHKATTEELKLKEVRKMSGYACDISDSPTVHKNVPNOVAKDPEKKLPLHVNTAGYE | 111 |
| Dd | : : : : : : : : : | |
| | 61 KKA-----EETLKRYGVKVKYKVNQSDVDVERFNVQVSDFEKIDAFIAN--AGATA | 112 |
| Oy | 112 NFPCEEDYPKANAKMKWKNLISLVYSOAPFKPLIKEGIKASVVLIGSMGAIIVNDPON | 171 |
| Dd | : : : : : : : : : | |
| | 113 NSGVVDGSASDMDHVIQVDLISGTAYCAKAVGAHFKEQG--HGSLVTATSMGSHVNPGE | 170 |
| Oy | 172 QVVYNNSKGVIIHAKTACEWAKNIRVNSIPGYTYGPTLNKVINGNEELYNRITSCI | 231 |
| Dd | : : : : : : : : : | |
| | 171 QTSYVNAKKGCHLARSLANEMDF--ARVNSISPXYIDTGLSDFIDEKTOEL---WRSMI | 226 |
| Oy | 232 PQGRMEPEKYTAGVLYLISESAASTTTGASLLVDGGFNS | 271 |
| Dd | : : : : : : : : : | |
| | 227 PMGRNGDAKELKGAYYVLVSD--ASSYTTGADIVIDGGYTT | 265 |
| | | |
| RESULT 6 | | |
| ID | ADJ98100 | |
| XX | ADJ98100 standard; protein; 283 AA. | |
| XX | ADJ98100; | |
| XX | | |
| XX | 06-MAY-2004 (first entry) | |
| XX | | |
| XX | Retoreductase (KRED) protein SEQ ID NO:42. | |

| | | |
|----|---|--|
| XX | 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester; | |
| KM | 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide; | |
| KM | 4-halo-3-ketobutyric acid ester; 3-hydroxybutyric acid ester; amide; | |
| KM | 4-nucleophile substituted 3-hydroxybutyric acid ester; amide; | |
| KM | ketoreductase; cofactor; cofactor regeneration system; nucleophile; | |
| KM | 4-nucleophile substituted-3-hydroxybutyric acid amide; | |
| KM | 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme. | |
| XX | | |
| OS | Synthetic. | |
| XX | | |
| PN | WO2004015132-A2. | |
| PD | 19-FEB-2004. | |
| XX | | |
| PF | 11-AUG-2003; 2003WO-US025263. | |
| PR | 09-AUG-2002; 2002US-0402436P. | |
| PR | 11-AUG-2003; 2003US-00402436. | |
| XX | | |
| PA | (CODE-) CODEXIS INC. | |
| PI | Davis SC, Grate JH, Gray DR, Gruber JM, Huisman GW, Ma SK; | |
| PI | Newman IM, Sheldon R, Wang LA; | |
| XX | | |
| DR | WPI: 2004-238752/22. | |
| XX | N-PSDB; ADJ98099. | |
| PT | Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3- | |
| PT | hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric | |
| PT | acid ester with halohydrin dehalogenase and cyanide. | |
| XX | | |
| PS | Example 9; SEQ ID NO 42; 168pp; English. | |
| XX | | |
| CC | The present invention describes a method (M1) for producing 4-cyano-3- | |
| CC | hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester | |
| CC | (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II) | |
| CC | or (III), and contacting (II) with a halohydrin dehalogenase and cyanide | |
| CC | under conditions sufficient to form a reaction mixture for converting | |
| CC | (II) to (I). Also described: (1) producing (M2) a 4-nucleophile | |
| CC | substituted 3-hydroxybutyric acid ester or amide from (II) or amide, | |
| CC | comprising: (a) providing (II), (III) or amide; and (b) contacting (II) | |
| CC | or amide with a halohydrin dehalogenase and a nucleophile under | |
| CC | conditions suitable to form a reaction mixture for converting (II) or | |
| CC | amide to a 4-nucleophile substituted-3-hydroxybutyric acid or amide; or | |
| CC | amide to a 4-nucleophile substituted-3-hydroxybutyric acid or amide; or | |
| CC | (c) contacting (III) or amide with a ketoreductase, a cofactor, a | |
| CC | cofactor regeneration system, a nucleophile and a halohydrin dehalogenase | |
| CC | to form a reaction mixture for converting (III) or amide to a 4- | |
| CC | nucleophile substituted-3-hydroxybutyric acid ester or amide; and (2) a | |
| CC | composition (IV) comprising halohydrin dehalogenase, a nucleophile, and | |
| CC | (I) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric | |
| CC | acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3- | |
| CC | ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile | |
| CC | substituted-3-hydroxybutyric acid ester or amide. (M1) and (M2) are | |
| CC | useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The | |
| CC | present sequence represents a ketoreductase (KRED), which is used in the | |
| CC | embodiment of the present invention. | |
| XX | | |
| SQ | Sequence 283 AA; | |
| QY | Query Match 26.9%; Score 381; DB 8; Length 283; | |
| QY | Best Local Similarity 36.1%; Pred. No. 1.9e-30; | |
| QY | Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8 | |
| Db | 8 FREDHLLVTGACGGALAEALIKGLAYGSDIALDIDDEKTAQAQAEYHKYATEBLK 67 | |
| Db | 29 FTKNGKVASITSSSGSIGYALAEFAQVADVAIWNSDATGKAEALAKKGV----- 82 | |
| QY | 68 EYPRKMSYACDIDSPTTHAKVPAQVAKPGKPLHLVNTAG--XCENPFCEDYPAKMAEK 125 | |
| Db | 83 ---KVAAYANVSSSDPAVKQTIEQOIKDGHLDIVVAN-AGIPWTKGAVITDDDDDHFDQ 138 | |
| QY | 126 MKVNLGLSLVYS---QAFAPRLIKEGIKGASVYLIGSMGSAIVDPONQVYNNMSKAG 181 | |

DB 139 VIDVDLKGVGVAKHAGHRYERREKEGKGA-LIFTSVSGHIVNIPOFQATYNAKAG 197
 QY 182 VIHLAKTACWAKNINRVNSLNPYIGPLTKRVINGNEELYNRMISGIPQORMSEPK 241
 DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTEISDFV---PQETQKMWSLVPLGRGGETAE 253
 QY 242 YIGAVLYLSSAASYTTGASLVDGGFT 270
 DB 254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 7

ADJ98118
 ID ADJ98118 standard; protein; 283 AA.

ADJ98118;
 DT 06-MAY-2004 (first entry)

XX Ketoreductase (KRED) protein SEQ ID NO:60.

XX 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
 XX 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide;
 KM 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
 KM ketoreductase; cofactor; cofactor regeneration system; nucleophile;
 KM 4-nucleophile substituted-3-hydroxybutyric acid amide;
 KM 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.

XX Synthetic.

XX WO2004015132-A2.

XX 19-FEB-2004.

XX 11-AUG-2003; 2003WO-US025263.

XX 09-AUG-2002; 2002US-0402436P.

XX 11-AUG-2003; 2003US-00402436.

XX (CODE-) CODEXIS INC.

PI Davis SC, Grate JH, Gray DR, Gruber JM, Huisman GW, Ma SK;
 PI Newman LM, Sheldon R, Wang LA;

XX WPI; 2004-238752/22.
 DR N-PSDB; ADJ98117.

PT Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-
 PT hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric
 PT acid ester with halohydrin dehalogenase and cyanide.

PS Disclosure; SEQ ID NO 60; 168pp; English.

XX The present invention describes a method (M1) for producing 4-cyano-3-
 CC hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester
 CC (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II)
 CC or (III), and contacting (II) with a halohydrin dehalogenase and cyanide
 CC under conditions sufficient to form a reaction mixture for converting
 CC (II) to (I). Also described: (1) producing (M2) a 4-nucleophile
 CC substituted 3-hydroxybutyric acid ester or amide from (II) or amide,
 CC comprising: (a) providing (II), (III) or amide; and (b) contacting (II)
 CC or amide with a halohydrin dehalogenase and a nucleophile under
 CC conditions suitable to form a reaction mixture for converting (II) or
 CC amide to a 4-nucleophile substituted-3-hydroxybutyric acid or amide; or
 CC (c) contacting (III) or amide with a ketoreductase, a cofactor, a
 CC cofactor regeneration system, a nucleophile and a halohydrin dehalogenase
 CC to form a reaction mixture for converting (III) or amide to a 4-
 CC nucleophile substituted-3-hydroxybutyric acid ester or amide; and (2) a
 CC composition (IV) comprising halohydrin dehalogenase, a nucleophile, and
 CC (I) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric
 CC acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3-
 CC ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile
 CC substituted-3-hydroxybutyric acid ester or amide. (M1) and (M2) are

CC useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The
 CC present sequence represents a ketoreductase (KRED), which is used in the
 CC exemplification of the present invention.

XX Sequence 283 AA;

Query Match 26.9%; Score 381; DB 8; Length 283;
 Best Local Similarity 36.1%; Pred. No. 1.9e-30;
 Matches 97; Conservative 52; Mismatches 98; Indels 22; Gaps 8;

QY 8 FRPDGHLITVAGCGALAEALIKGLIANGSDIALDIDOEKTAQKQAEHYHATBELTK 67
 DB 29 FKLNGKVASITGSSSGIGYALAEAPQAVGADVAIWMNSQDRTGAELAKKTYG----- 82
 QY 68 EVPKMSYACDISDPTVHKVFAQVAKDFGLPLHLVNTAG--YCENPCEBEPYAKNAEK 125
 DB 83 ---KTKAVKANVSSSDAVKQITIEQIDRFGHLDIVAN-AGIPMTKAYIIODDDKHDQ 138
 QY 126 MKRVNLGSLYVS---QAFKPLIKEGIKASVVLGSMGAIIVDPQNVVYMSKAG 181
 DB 139 VIDVDLKGVGVAKHAGHRYERREKEGKGA-LIFTSVSGHIVNIPOFQATYNAKAG 197
 QY 182 VIHLAKTACWAKNINRVNSLNPYIGPLTKRVINGNEELYNRMISGIPQORMSEPK 241
 DB 198 VRHFAKSLAVEFAPF-ARVNSVSPGYINTEISDFV---PQETQKMWSLVPLGRGGETAE 253
 QY 242 YIGAVLYLSSAASYTTGASLVDGGFT 270
 DB 254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 8

ADJ98098
 ID ADJ98098 standard; protein; 283 AA.

ADJ98098;
 DT 06-MAY-2004 (first entry)

XX Ketoreductase (KRED) protein SEQ ID NO:40.

XX 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
 KM 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide;
 KM 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
 KM ketoreductase; cofactor; cofactor regeneration system; nucleophile;
 KM 4-nucleophile substituted-3-hydroxybutyric acid amide;
 KM 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.

XX Synthetic.

XX WO2004015132-A2.

XX 19-FEB-2004.

XX 11-AUG-2003; 2003WO-US025263.

XX 09-AUG-2002; 2002US-0402436P.

XX 11-AUG-2003; 2003US-00402436.

XX (CODE-) CODEXIS INC.

PI Davis SC, Grate JH, Gray DR, Gruber JM, Huisman GW, Ma SK;
 PI Newman LM, Sheldon R, Wang LA;

XX WPI; 2004-238752/22.
 DR N-PSDB; ADJ98097.

PT Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-
 PT hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric
 PT acid ester with halohydrin dehalogenase and cyanide.

PS Disclosure; SEQ ID NO 40; 168pp; English.

XX 31-JAN-2001; 2001WO-US003186.
XX
XX
XX 31-JAN-2000; 2000US-00494921.
XX
XX 08-DEC-2000; 2000US-00734237.
XX
XX (BIOC-) BIOCATALYTICS INC.
XX
XX Roszell DJ, Bui P, Hua L;
XX
XX WPI; 2001-483235/52.
XX
XX N-PSDB; AAH74597.
XX
XX
XX Designing synthetic nucleic acid sequences for improved amplification,
XX
XX PT expression in host cell, by comparing free energy of folding of a
XX
XX PT starting polynucleotide and a modified polynucleotide having a codon
XX
XX PT replacement.
XX
XX
XX Claim 6; Page 103; 117pp; English.
XX
XX
XX The present sequence is a NADPH-dependent carbonyl reductase. The
XX
XX CC polynucleotide was modified using the method of the invention. The
XX
XX CC specification describes a method for designing a synthetic
XX
XX CC polynucleotide. The method comprises providing a starting polynucleotide,
XX
XX CC determining the predicted free energy of folding per base of the
XX
XX CC polynucleotide, modifying the polynucleotide by replacing a codon with a
XX
XX CC different codon to provide a modified polynucleotide, determining free
XX
XX CC energy of folding per base of the modified polynucleotide, and comparing
XX
XX CC this with that of the original polynucleotide. The method is useful for
XX
XX CC developing nucleic acid sequences that enhance expression of the encoded
XX
XX CC protein in a heterologous host. The design and preparation of the
XX
XX CC synthetic genes are used in application of gene shuffling, directed
XX
XX CC evolution and molecular breeding methods. The method allows expression of
XX
XX CC genes from various organisms such as mammals, plants, yeast, fungi and
XX
XX CC bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic
XX
XX CC hosts at commercially viable levels, in particular proteins with low
XX
XX CC yield such as methionine gamma-lyase from *P. putida*
XX
XX
XX Sequence 283 AA;
XX
XX
XX Query Match 26.7%; Score 378; DB 4; Length 283;
XX
XX Best Local Similarity 36.8%; Pred. No. 3.9e-30;
XX
XX Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;
XX
XX
XX 8 FRPDGHTITVAGCGIAEALIKGLAYGSDIALDIDOEKTAQAQEHYKATBELK 67
XX
XX DB 29 FKUNGKVASITGSSGIGYALAEAFQVGDVAIWNVSHDATGKAELAKKYGV----- 82
XX
XX QY 68 EVPKMGSYACDISDPTVHVFAQVAKDPGKLPPLVNTAG--YCENFPCEDPYAKNAEK 125
XX
XX DB 83 ---KVKAAYKANVSSSDAVKQTFIEQIQIDFGHLDIVAN--AGIPWTKGAYIDQDDDKHFDQ 138
XX
XX QY 126 MVKNNLLGSLYVS---QAFAPKPLIKEGIGKASVVLIGSMGALVNDPQNVVYNNMSKAG 181
XX
XX DB 139 VVDVDLGQGVAAHAGHFERERKEGKKA-LVFTASMSGHIVNVPQFATYNAKAG 197
XX
XX QY 182 VIHLAKTLACEMAKNIRVNSLNPGYTGPLTKVINGNEELVNRWISGIPQOMSEPK 241
XX
XX DB 198 VRHRAKSLAVEFAF-ARVSVSROGYINTETSDV---PQETQKMSVLVLGSGEYTAB 253
XX
XX QY 242 YIGAVLYLSESAASYTTGASLVDGGEFT 270
XX
XX DB 254 LVGAYLFLASD-AGSYATGTDIIVDGGYT 281
XX
XX
XX RESULT 11
XX
XX ID ADJ98066 standard; protein; 283 AA.
XX
XX AC ADJ98066;
XX
XX DT 06-MAY-2004 (first entry)
XX

DE Ketoreductase 4 (KRED CR2-4) protein SEQ ID NO:8.
XX
XX
XX 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
XX
XX KM 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide;
XX
XX KM 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
XX
XX KM ketoreductase; cofactor; cofactor regeneration system; nucleophile;
XX
XX KM 4-nucleophile substituted 3-hydroxybutyric acid amide;
XX
XX KM 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.
XX
XX
XX *Candida magnoliae*.
XX
XX
XX WO2004015132-A2.
XX
XX
XX 19-FEB-2004.
XX
XX
XX 11-AUG-2003; 2003WO-US025263.
XX
XX
XX 09-AUG-2002; 2002US-0402436P.
XX
XX PR 11-AUG-2003; 2003US-00402436.
XX
XX
XX (CODE-) CODEXIS INC.
XX
XX
XX Davis SC, Grate JH, Gray DR, Gruber JM, Huisman GW, Ma SK;
XX
XX PI Newman LM, Sheldon R, Wang LA;
XX
XX WPI; 2004-238752/22.
XX
XX DR N-PSDB; ADJ98065.
XX
XX
XX Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-
XX
XX PT hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric
XX
XX PT acid ester with halohydrin dehalogenase and cyanide.
XX
XX
XX Example 1, SEQ ID NO 8; 16pp; English.
XX
XX
XX The present invention describes a method (M1) for producing 4-cyano-3-
XX
XX CC hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester
XX
XX CC (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II)
XX
XX CC or (III), and contacting (II) with a halohydrin dehalogenase and cyanide
XX
XX CC under conditions sufficient to form a reaction mixture for converting
XX
XX CC (II) to (I). Also described: (1) producing (M2) a 4-nucleophile
XX
XX CC substituted 3-hydroxybutyric acid ester or amide from (II) or amide,
XX
XX CC comprising: (a) providing (II), (III) or amide; and (b) contacting (II)
XX
XX CC or amide with a halohydrin dehalogenase and a nucleophile under
XX
XX CC conditions suitable to form a reaction mixture for converting (II) or
XX
XX CC amide to a 4-nucleophile substituted 3-hydroxybutyric acid or amide; or
XX
XX CC (c) contacting (III) or amide with a ketoreductase, a cofactor, a
XX
XX CC cofactor regeneration system, a nucleophile and a halohydrin dehalogenase
XX
XX CC to form a reaction mixture for converting (III) or amide to a 4-
XX
XX CC nucleophile substituted 3-hydroxybutyric acid ester or amide; and (2) a
XX
XX CC composition (IV) comprising halohydrin dehalogenase, a nucleophile, and
XX
XX CC (1) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric
XX
XX CC acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3-
XX
XX CC ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile
XX
XX CC substituted 3-hydroxybutyric acid ester or amide. (M1) and (M2) are
XX
XX CC useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The
XX
XX CC present sequence represents a ketoreductase (KRED), which is used in the
XX
XX CC exemplification of the present invention.
XX
XX
XX Sequence 283 AA;
XX
XX
XX Query Match 26.7%; Score 378; DB 8; Length 283;
XX
XX Best Local Similarity 36.8%; Pred. No. 3.9e-30;
XX
XX Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;
XX
XX
XX 8 FRPDGHTITVAGCGIAEALIKGLAYGSDIALDIDOEKTAQAQEHYKATBELK 67
XX
XX DB 29 FKUNGKVASITGSSGIGYALAEAFQVGDVAIWNVSHDATGKAELAKKYGV----- 82
XX
XX QY 68 EVPKMGSYACDISDPTVHVFAQVAKDPGKLPPLVNTAG--YCENFPCEDPYAKNAEK 125
XX
XX DB 83 ---KVKAAYKANVSSSDAVKQTFIEQIQIDFGHLDIVAN--AGIPWTKGAYIDQDDDKHFDQ 138
XX
XX QY 126 MVKNNLLGSLYVS---QAFAPKPLIKEGIGKASVVLIGSMGALVNDPQNVVYNNMSKAG 181
XX

| | | | |
|-----------|---|--|-----|
| D6 | | 139 VVDVILKGVYAKNAGRFRFRFEKEGKKGA-LVFATMSMGIIVVPFOATTYNAKGG | 197 |
| Oy | | 182 VHLKLKTACAEWAKNIIRVNSLNPGYTGLPKNVINGNEELYNRWISGIPPOORMEPEKE | 241 |
| D6 | | 198 VHFPMKSLAVEPAPF-ARVNVSYPGITINTEISFPV---PQETONKMMSLVPLGRGETAE | 253 |
| Oy | | 242 YIGAVLYLLSESASYYTTGASLLVDGFGFT 270 | |
| D6 | | 254 LVGAVTLFLASD-AGSVATGTDTIIVDGCT 281 | |
| Db | | | |
| RESULT 12 | | | |
| ADJ98060 | | | |
| ID | ADJ98060 | standard; protein; 283 AA. | |
| XX | | | |
| XX | ADJ98060; | | |
| AC | | | |
| XX | | | |
| DT | 06-MAY-2004 | (first entry) | |
| XX | | | |
| DE | Ketoreductase 1 (KRED CR2-5) | protein SEQ ID NO:2. | |
| XX | | | |
| KM | 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester; | | |
| KM | 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide; | | |
| KM | 4-nucleophile substituted 3-hydroxybutyric acid ester; amide; | | |
| KM | ketoerectase; cofactor; cofactor regeneration system; nucleophile; | | |
| KM | 4-nucleophile substituted-3-hydroxybutyric acid amide; | | |
| KM | 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme. | | |
| OS | Candida magnoliae. | | |
| XX | | | |
| PN | WO2004015132-A2. | | |
| XX | | | |
| PD | 19-FEB-2004. | | |
| XX | | | |
| Pf | 11-AUG-2003; 2003WC-US025263. | | |
| PR | 09-AUG-2002; 2002US-0402436P. | | |
| PR | 11-AUG-2003; 2003US-00402436. | | |
| XX | | | |
| PA | (CODE-) CODEXIS INC. | | |
| P1 | David SC, Grate JH, Gray DR, Gruber JM, Huismann GW, Ma SK; | | |
| P1 | Newman IM, Sheldon R, Wang LA; | | |
| DR | MP1: 2004-238752/22. | | |
| DR | N-PADB; ADJ98059. | | |
| XX | | | |
| PT | Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3- | | |
| PT | hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric | | |
| PT | acid ester with halohydrin dehalogenase and cyanide. | | |
| XX | | | |
| XX | Example 1; SEQ ID NO 2; 168bp; English. | | |
| XX | | | |
| CC | The present invention describes a method (M1) for producing 4-cyano-3- | | |
| CC | hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester | | |
| CC | (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II) | | |
| CC | or (III), and contacting (II) with a halohydrin dehalogenase and cyanide | | |
| CC | under conditions sufficient to form a reaction mixture for converting | | |
| CC | (II) to (I). Also described: (1) producing (M2) a 4-nucleophile | | |
| CC | substituted 3-hydroxybutyric acid ester or amide from (II) or amide, | | |
| CC | comprising: (a) providing (II), (III) or amide; and (b) contacting (II) | | |
| CC | or amide with a halohydrin dehalogenase and a nucleophile under | | |
| CC | conditions suitable to form a reaction mixture for converting (II) or | | |
| CC | amide to a 4-nucleophile substituted-3-hydroxybutyric acid or amide; or | | |
| CC | (c) contacting (III) or amide with a ketoerectase, a cofactor, a | | |
| CC | cofactor regeneration system, a nucleophile and a halohydrin dehalogenase | | |
| CC | to form a reaction mixture for converting (III) or amide to a 4- | | |
| CC | nucleophile substituted-3-hydroxybutyric acid ester or amide; and (2) a | | |
| CC | composition (IV) comprising halohydrin dehalogenase, a nucleophile, and | | |
| CC | (1) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric | | |
| CC | acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3- | | |
| CC | ketoerectase acid ester. (M2) is useful for producing a 4-nucleophile | | |

CC substituted-3-hydroxybutyric acid ester or amide; (M1) and (M2) are
CC useful for producing 4-substituted 3-hydroxybutyric acid derivative. The
CC present sequence represents a ketoreductase (KRED), which is used in the
XX exemplification of the present invention.

Seq Sequence 263 AA;

| | Query Match | 26.7% | Score 378; | DB 8; | Length 263; |
|--|-----------------------|--------|--------------------|-----------------|-------------------|
| | Best Local Similarity | 36.8%; | Pred. No. 3.9e-30; | | |
| | Matches | 99; | Conservative 48; | Mismatches 100; | Indels 22; Gaps 8 |

OY 8 FRPDGHLTVVGACGALAEALIKGLAYGPSDALDIDPOEKTPAAQAEHYKYATEELKLK 67
| : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 29 FLNKKVASTIGSSSGICGYALLAQAQGVADVAIYNSHDATGKAELAKKTGV----- 82
68 EYPKNGSYACDISDSDTYHKVPYAQAFGKLP LHLVNTAG--YCENPFCEDDYPAKNAEK 125
OY 83 ---KTKAVKANVSSSDAVKMQTIEQIDKFHGHDIVVAN-AGIPWTKGAVIIDODDDGHFDQ 138
126 MYKNLLLSLYS---QAFAKPLIKEGIKASVYLIGSMGSAIYNNDPQNQVYNNMSKG 181
Db 139 VVDVVLKGGVYAKHAGHHFEREPKEGKGA-LVFATSMSGHIVNVPPFQATYNAAKG 197
OY 182 VTILAKTLACEWAKTNIRVNSLNPQTYTGPLTKNYINGNEELYNRNISGIPOORMSEPKE 241
198 VNHFKSLAVEPAFP-ARVNSVSPGIYNTETDFV---PQETONKWSLVPLGRGETAE 253
OY 242 YIGAVLYLSESASA SYTGASILLVDGGFT 270
| : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 254 LVGAYLFLASD-AGSYATGTDLIVDGGYT 281
RESULT 13
ADJ98064
ID ADJ98064 strand; protein; 283 AA.
AC ADJ98064;
XX
DT 06-MAY-2004 (first entry)
XX
DE Ketoreductase 3 (KRED CRI-3) protein SEQ ID NO:6.
XX
KW 4-cyano-3-hydroxybutyric acid ester; 4-halo-3-hydroxybutyric acid ester;
KW 4-halo-3-ketobutyric acid ester; halohydrin dehalogenase; cyanide;
KW 4-nucleophile substituted 3-hydroxybutyric acid ester; amide;
KW ketoreductase; cofactor; cofactor regeneration system; nucleophile;
KW 4-nucleophile substituted-3-hydroxybutyric acid amide;
KW 4-substituted 3-hydroxybutyric acid derivative; KRED; enzyme.
XX
OS Candida magnoliae.
XX
PN WO2004015132-A2.
XX
PD 19-FEB-2004.
XX
PF 11-AUG-2003; 2003WO-US025263.
XX
PR 09-AUG-2002; 2002US-0402436P.
XX
PR 11-AUG-2003; 2003US-00402436.
XX
PA (CODE-) CODEXIS INC.
XX
PI Davis SC, Grace JH, Gray DR, Gruber JM, Huismann GW, Ma SK,
PI Newman LM, Sheldon R, Wang LA;
XX
DR WFI, 2004-238752/22.
XX
DR N-PDB; ADJ98063.
PT Producing 4-cyano-3-hydroxybutyric acid ester from 4-halo-3-
PT hydroxybutyric acid ester, comprises contacting 4-halo-3-hydroxybutyric
PT acid ester with halohydrin dehalogenase and cyanide.
XX
XX Example 1; SEQ ID NO 6; 16bp; English.

XX The present invention describes a method (M1) for producing 4-cyano-3-hydroxybutyric acid ester (I) from a 4-halo-3-hydroxybutyric acid ester (II) or 4-halo-3-ketobutyric acid ester (III), comprising providing (II) or (III), and contacting (II) with a halohydrix dehalogenase and cyanide under conditions sufficient to form a reaction mixture for converting (II) to (I). Also described: (1) producing (M2) a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide from (II) or amide, comprising: (a) providing (II), (III) or amide; and (b) contacting (II) or amide with a halohydrix dehalogenase and a nucleophile under conditions suitable to form a reaction mixture for converting (II) or amide to a 4-nucleophile substituted 3-hydroxybutyric acid or amide; or (c) contacting (III) or amide with a ketoreductase, a cofactor, a cofactor regeneration system, a nucleophile and a halohydrix dehalogenase to form a reaction mixture for converting (III) or amide to a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide; and (2) a composition (IV) comprising halohydrix dehalogenase, a nucleophile, and (1) or amide. (M1) is useful for producing a 4-cyano-3-hydroxybutyric acid ester from a 4-halo-3-hydroxybutyric acid ester or 4-halo-3-ketobutyric acid ester. (M2) is useful for producing a 4-nucleophile substituted 3-hydroxybutyric acid ester or amide. (M1) and (M2) are useful for producing 4-substituted 3-hydroxybutyric acid derivatives. The present sequence represents a ketoreductase (KRED), which is used in the exemplification of the present invention.

CC Sequence 283 AA;

Query Match 26.7%; Score 378; DB 8; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.9e-30;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLTVTGACGGLAEALIKGLAYGSDIALDIDOEKTAQAQEHYKATBELKUK 67
DB 29 FDLNKGVALVTVGAGSGIGQRIAMGLAGAGAVALLDRRTDGLAQTAQFIARAG----- 82
QY 68 EYPMKGSYACDISDSDYTHKVPAAQVAFGLPLHLVNTAG--YCNFCEDEYPAKNAEK 125
DB 83 ---KVAKAVANSSDAVAKQTIEQOI KDFGHLDIVAN--AGIPWKGAVIDDDGHFQO 138
QY 126 MYKVMILSLVYS---QAFAPLKEGIGKASVVLGSGAVLVNDPONOVVNMSSKAG 181
DB 139 VVDVVLKGVGVYAKAGARFRFRFEKGGKGA-LFTTASMSGHIVAVPOFATTNAKAG 197
QY 182 VIHLAKTACEMAKYINRVNSLMPGIVPLTKVINGNEBELYNRMISGIPQORMSEPKK 241
DB 198 VTHFAKSLAVEPAPF-ARVNSVSPGYINTEISDFV---PQETQNMKWSLVPLGRGETAB 253
QY 242 YIGAVLYLSSASAYTTGASLLVVDGFT 270
DB 254 LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281

RESULT 14

ADS22515 standard; protein; 253 AA.

ADS22515;

02-DEC-2004 (first entry)

Bacterial polypeptide #11548.

Recombinant DNA construct; transformed plant; improved plant property;
cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
pathogen tolerance; pest tolerance; plant disease resistance;
cell cycle pathway modification; plant growth regulator;
homologous recombination; seed oil yield; protein yield; carbohydrate;
nitrogen; photosynthesis; lignin; galactomannan;
bacterial polypeptide.

Bacteria.

US2003233675-A1.

XX 18-DEC-2003.
PD 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
DR WPI; 2004-061375/06.
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
PS Claim 1, SEQ ID NO 11548; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

CC Sequence 253 AA;

Query Match 26.6%; Score 376; DB 8; Length 253;
Best Local Similarity 35.8%; Pred. No. 5.3e-30;
Matches 96; Conservative 43; Mismatches 109; Indels 20; Gaps 7;

QY 8 FRPDGHLTVTGACGGLAEALIKGLAYGSDIALDIDOEKTAQAQEHYKATBELKUK 67
DB 3 FDLNKGVALVTVGAGSGIGQRIAMGLAGAGAVALLDRRTDGLAQTAQFIARAG----- 57
QY 68 EYPMKGSYACDISDSDYTHKVPAAQVAFGLPLHLVNTAGYCNFCEDEYPAKNAEKV 127
DB 58 ---KSIQIAADVTSKQALTDVAVARQAEIAGLSL--AVNAAGIANPAAEMSESOFTWM 113
QY 128 KVMILSLVYSQAAPKPLKEGIGKASVVLGSGAVLVNDPONOVVNMSSKAGYTHLAK 187
DB 114 DINLKGVLSCQAEENAMTKNG-KGA-IVNTASMSGVIVNGLMQCHYNASAGVITHSK 171
QY 188 TLACEMAKYINRVNSLMPGIVPLTKVINGNEBELYNR---WISGIPQORMSEPKK 244
DB 172 SMAEMWVGRIKRVMTISPGYATPM-----WTRPMVHQTFLPBGQTMMAGADVAVG 226
QY 245 AVLVLSSASAYTTGASLLVVDGFTSW 272
DB 227 PAIFLLSD-AASFVTVGVDLVDGGFCW 253

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:21:36 ; Search time 43 Seconds
(without alignments)
472.199 Million cell updates/sec

Title: US-10-720-018-2
Perfect score: 1415
Sequence: 1 MTDVPTFRFGHLLTVTGA.....SAASYTTGASLLVDGFTSW 272

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 686 | 48.5 | 231 | US-08-336-198C-7 | Sequence 7, Appli |
| 2 | 686 | 48.5 | 231 | US-09-184-965-7 | Sequence 7, Appli |
| 3 | 586 | 41.4 | 282 | US-09-248-796A-14132 | Sequence 14132, A |
| 4 | 378 | 26.7 | 283 | US-09-367-012-1 | Sequence 1, Appli |
| 5 | 378 | 26.7 | 283 | US-09-777-157A-1 | Sequence 10, Appli |
| 6 | 378 | 26.7 | 283 | US-09-734-237B-70 | Sequence 70, Appli |
| 7 | 372.5 | 26.3 | 308 | US-09-248-796A-16593 | Sequence 16593, A |
| 8 | 344 | 24.3 | 316 | US-09-489-039A-12990 | Sequence 12990, A |
| 9 | 340.5 | 24.1 | 257 | US-09-134-001C-3562 | Sequence 3562, Ap |
| 10 | 332.5 | 22.8 | 256 | US-09-252-991A-30474 | Sequence 30474, A |
| 11 | 321.5 | 22.7 | 262 | US-09-363-189B-6 | Sequence 6, Appli |
| 12 | 312.5 | 22.1 | 292 | US-09-468-738A-2 | Sequence 2, Appli |
| 13 | 312.5 | 22.1 | 292 | US-09-940-019-2 | Sequence 2, Appli |
| 14 | 312.5 | 22.1 | 292 | US-09-940-037A-2 | Sequence 2, Appli |
| 15 | 312.5 | 22.1 | 296 | US-09-468-738A-23 | Sequence 23, Appli |
| 16 | 312.5 | 22.1 | 296 | US-09-940-019-23 | Sequence 23, Appli |
| 17 | 312.5 | 22.1 | 296 | US-09-940-037A-23 | Sequence 23, Appli |
| 18 | 307 | 21.7 | 251 | US-09-922-501-16 | Sequence 16, Appli |
| 19 | 302 | 21.3 | 275 | US-09-107-532A-4805 | Sequence 4805, Ap |
| 20 | 296 | 20.9 | 249 | US-09-902-540-1561 | Sequence 14561, A |
| 21 | 290 | 20.5 | 258 | US-09-504-358-12 | Sequence 12, Appli |
| 22 | 290 | 20.5 | 258 | US-09-954-314-12 | Sequence 12, Appli |
| 23 | 290 | 20.5 | 258 | US-10-230-562-12 | Sequence 12, Appli |
| 24 | 289 | 20.4 | 256 | US-08-594-808B-7 | Sequence 7, Appli |
| 25 | 282 | 19.9 | 290 | US-09-248-796A-16592 | Sequence 16592, A |
| 26 | 279 | 19.7 | 303 | US-09-002-298-1 | Sequence 1, Appli |
| 27 | 279 | 19.7 | 303 | US-09-481-277-1 | Sequence 1, Appli |

| | | | | | |
|----|-------|------|-----|----------------------|--------------------|
| 28 | 279 | 19.7 | 322 | US-09-634-955B-11 | Sequence 11, Appli |
| 29 | 279 | 19.7 | 322 | US-09-816-760-11 | Sequence 11, Appli |
| 30 | 279 | 19.7 | 322 | US-09-838-561-11 | Sequence 11, Appli |
| 31 | 273.5 | 19.3 | 290 | US-09-248-796A-16757 | Sequence 16757, A |
| 32 | 270.5 | 19.1 | 269 | US-09-134-000C-4929 | Sequence 4929, Ap |
| 33 | 268 | 18.9 | 262 | US-09-328-352-5742 | Sequence 5742, Ap |
| 34 | 265.5 | 18.8 | 256 | US-09-543-681A-6693 | Sequence 6693, Ap |
| 35 | 265.5 | 18.8 | 276 | US-09-270-767-41628 | Sequence 41628, A |
| 36 | 264.5 | 18.7 | 274 | US-09-134-001C-4431 | Sequence 4431, Ap |
| 37 | 264.5 | 18.7 | 306 | US-09-710-279-382 | Sequence 382, Ap |
| 38 | 263 | 18.6 | 261 | US-09-328-352-7731 | Sequence 7731, Ap |
| 39 | 262 | 18.5 | 303 | US-09-634-955B-25 | Sequence 25, Appli |
| 40 | 262 | 18.5 | 303 | US-09-816-760-13 | Sequence 13, Appli |
| 41 | 262 | 18.5 | 303 | US-09-838-561-13 | Sequence 13, Appli |
| 42 | 261 | 18.4 | 266 | US-09-107-433-3795 | Sequence 3795, Ap |
| 43 | 260.5 | 18.4 | 254 | US-09-978-758-2 | Sequence 2, Appli |
| 44 | 259.5 | 18.3 | 259 | US-09-489-039A-11936 | Sequence 11936, A |
| 45 | 257.5 | 18.2 | 308 | US-09-489-039A-7802 | Sequence 7802, Ap |

ALIGNMENTS

RESULT 1
US-08-336-198C-7
; Sequence 7, Application US/08336198C
; Patent No. 5866382
; GENERAL INFORMATION:
; APPLICANT: Hallborn, Johan
; APPLICANT: Penttila, Merja
; APPLICANT: Ojamo, Heikki
; APPLICANT: Keranen, Sirkka
; APPLICANT: Hahn-Hagerdal, Barbel
; APPLICANT: Waldfriedson, Mats
; APPLICANT: Alaksinen, Ulla
; TITLE OF INVENTION: XYLOSE UTILIZATION BY RECOMBINANT YEASTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,198C
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-336-198C-7
Query Match 48.5% Score 686; DB 2; Length 231;
Best Local Similarity 55.6%; Pred. No. 2.2e-67;
Matches 130; Conservative 49; Mismatches 49; Indels 6; Gaps 3;
QY 42 LDIDQKTAQAQAEYHKYATEELKKEVPRWG---SYACDISDSPTVAKFAQVAKQFGK 98

Db 1 MNLRTSAAK--EVLGGEETLKGHEHSAIGQVSAMSCNIGDAEADATFSSINEHKG 58
Qy 99 LPLHLVNTAGYCNFPCEDYPAKNAEKVKNLIGSLVYSGAPFKPLIKESIKASVLI 158
Db 59 IADLLINTAGYCNFPATYPAKNAESIMKVNGLGSFYVSQSPARPLIÖNNLRG-STILI 117
Qy 159 GMSGAIYVNDPÖQÖVMYNSKAGVIHLAKTLACEMAKYINRVNSLNPGYIYGPLTKVYN 218
Db 118 GMSGAIYVNDPÖQÖVMYNSKAGVIHLVRSIACEMAKYINRVNLSPOYILTLPLRVNIS 177
Qy 219 GNEELYNRMISGIPQÖRMSEPKYIGAVLYLSESASVTTGASLVDGFTSW 272
Db 178 GHEMKEMESKIPMKRMAEPKEFVGSILYLASETASVTTGHNLVVDGGYECW 231
RESULT 2
US-09-184-965-7
; Sequence 7, Application US/09184965
; Patent No. 6582944
; GENERAL INFORMATION:
; APPLICANT: Hallborn, Johan
; APPLICANT: Penttilä, Merja
; APPLICANT: Ojamo, Heikki
; APPLICANT: Keränen, Sirkka
; APPLICANT: Hahn-Hagerdal, Barbel
; APPLICANT: Walderidsson, Mats
; APPLICANT: Airaksinen, Ulla
; TITLE OF INVENTION: Xylose UTILIZATION BY RECOMBINANT YEASTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 N. Washington St.
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22046-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,965
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,198
; FILING DATE: 03-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-184-965-7
Query Match 48.5%; Score 686; DB 4; Length 231;
Best Local Similarity 55.6%; Pred. No. 2.2e-67;
Matches 130; Conservative 49; Mismatches 49; Indels 6; Gaps 3;
Qy 42 LIDDEKTAQAQAEHYKVAETELKKEVPMKG---SYACDISDPTVHKVPAQVAKDFK 98
Db 1 MNLRTSAAK--EVLGGEETLKGHEHSAIGQVSAMSCNIGDAEADATFSSINEHKG 58
Qy 99 LPLHLVNTAGYCNFPCEDYPAKNAEKVKNLIGSLVYSGAPFKPLIKESIKASVLI 158

Db 59 IADLLINTAGYCNFPATYPAKNAESIMKVNGLGSFYVSQSPARPLIÖNNLRG-STILI 117
Qy 159 GMSGAIYVNDPÖQÖVMYNSKAGVIHLAKTLACEMAKYINRVNSLNPGYIYGPLTKVYN 218
Db 118 GMSGAIYVNDPÖQÖVMYNSKAGVIHLVRSIACEMAKYINRVNLSPOYILTLPLRVNIS 177
Qy 219 GNEELYNRMISGIPQÖRMSEPKYIGAVLYLSESASVTTGASLVDGFTSW 272
Db 178 GHEMKEMESKIPMKRMAEPKEFVGSILYLASETASVTTGHNLVVDGGYECW 231
RESULT 3
US-09-248-796A-14132
; Sequence 14132, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14132
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14132
Query Match 41.4%; Score 586; DB 4; Length 282;
Best Local Similarity 54.6%; Pred. No. 3.5e-56;
Matches 113; Conservative 42; Mismatches 48; Indels 4; Gaps 2;
Qy 5 IPTFRFDGHLTIYTGACGGLAELIKGLLAGSDIALDIDOEKTAQAQAEHYKVAETEL 64
Db 73 VPSFRIDGKLVILVGGSGGLAAVSRALLAKGADVALLVDMNLERTQOARADVLMGGEOM 132
Qy 65 KLR--EYPMKGSYACDISDSDTYHKVPAQVAKFGKLPVHLVNTAGYCNFPCEDYPAK 121
Db 133 KGKYESPIGQVSAMSCNIGDAEADLTFKALNEHKGKISSVLVNTAGYAEVPAEYPAK 192
Qy 122 NAEKVVKNLIGSLVYSGAPFKPLIKESIKASVLI GMSGAIYVNDPÖQÖVMYNSKAG 181
Db 193 NAEKVVKNLIGSLVYSGAPFKPLIÖNNMTG-STILIGSGGAIYVNDPÖQÖVMYNSKAG 251
Qy 182 VIHLAKTLACEMAKYINRVNSLNPGYI 208
Db 252 VIHLAKTLACEMAKYINRVNLSPOYI 278
RESULT 4
US-09-367-012-1
; Sequence 1, Application US/09367012
; Patent No. 6218156
; GENERAL INFORMATION:
; APPLICANT: Yasonara, Yoshihiko
; APPLICANT: Kizaki, No. 6218156Iyuki
; APPLICANT: Hasegawa, Junzo
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Kataoka, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; APPLICANT: Kita, Keiko
; TITLE OF INVENTION: No. 6218156el carbonyl reductase, gene coding same, and method
; FILE REFERENCE: S72030PCUS
; CURRENT APPLICATION NUMBER: US/09/367,012

CURRENT FILING DATE: 1999-11-24
; EARLIER APPLICATION NUMBER: PCT/JP97/03051
; EARLIER FILING DATE: 1997-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Candida magnoliae
US-09-367-012-1

Query Match 26.7%; Score 378; DB 3; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.7e-33;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLITVAGCGLAELIKGLAYGSDIALDIDOEKTAQAQAEYHKATEELK 67
DB 29 FRLNKGKVASITGSSGIGALAEAFQVADVAIWNSHDAGKALAKKYG----- 82
QY 68 EYPRKGSYACDISDSDYHKKVFAQVAKDFGLPLHLVNTAG--YCENPCECDYPAKNAEK 125
DB 83 ---KVKAYKANVSSSDAVKQTEBOQIKDFGHLIDIVAN-AGIPWTKGAYIDDDDKHPDQ 138
QY 126 MVKNVNLGSLVYS---QAFAPLKEGKGSVVLIGSMGATVNDPQNVVYNNMSKAG 181
DB 139 VVDVDLKGVYAKAGHFRERFEKEGKGA-LVFTASMSGHIVNVQFOATYNAAKAG 197
QY 182 VIHLAKTACEMAKYNIRVNSLNPGYIYGLTKVINGNEELYNRMISGIPQORMSEPK 241
DB 198 VRHFAKSLAVEFAPE-ARVNSVSPGYINTBISDFV---PQETQKNKMSLVPLGRGETAE 253
QY 242 YIGAVLYLSSESASVTTGASLVLDGFT 270
DB 254 LVGAVLFLASD-AGSVATGTDIIVDGGYT 281

RESULT 5

US-09-777-157A-1
; Sequence 1, Application US/09777157A
; Patent No. 6448052
; GENERAL INFORMATION:
; APPLICANT: Yasohara, Yoshihiko
; APPLICANT: Kizaki, No. 6448052yuki
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Katoko, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; APPLICANT: Kita, Keiko
; TITLE OF INVENTION: Carbonyl Reductase Enzyme and Methods for its Use
; FILE REFERENCE: 068383.0110
; CURRENT APPLICATION NUMBER: US/09/777,157A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/367,012
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 1
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Candida magnoliae
US-09-777-157A-1

Query Match 26.7%; Score 378; DB 4; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.7e-33;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLITVAGCGLAELIKGLAYGSDIALDIDOEKTAQAQAEYHKATEELK 67
DB 29 FRLNKGKVASITGSSGIGALAEAFQVADVAIWNSHDAGKALAKKYG----- 82
QY 68 EYPRKGSYACDISDSDYHKKVFAQVAKDFGLPLHLVNTAG--YCENPCECDYPAKNAEK 125

DB 83 ---KVKAYKANVSSSDAVKQTEBOQIKDFGHLIDIVAN-AGIPWTKGAYIDDDDKHPDQ 138
QY 126 MVKNVNLGSLVYS---QAFAPLKEGKGSVVLIGSMGATVNDPQNVVYNNMSKAG 181
DB 139 VVDVDLKGVYAKAGHFRERFEKEGKGA-LVFTASMSGHIVNVQFOATYNAAKAG 197
QY 182 VIHLAKTACEMAKYNIRVNSLNPGYIYGLTKVINGNEELYNRMISGIPQORMSEPK 241
DB 198 VRHFAKSLAVEFAPE-ARVNSVSPGYINTBISDFV---PQETQKNKMSLVPLGRGETAE 253
QY 242 YIGAVLYLSSESASVTTGASLVLDGFT 270
DB 254 LVGAVLFLASD-AGSVATGTDIIVDGGYT 281

RESULT 6

US-09-734-237B-70
; Sequence 70, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzelli, J. David
; APPLICANT: But, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 70
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Candida magnoliae
US-09-734-237B-70

Query Match 26.7%; Score 378; DB 4; Length 283;
Best Local Similarity 36.8%; Pred. No. 3.7e-33;
Matches 99; Conservative 48; Mismatches 100; Indels 22; Gaps 8;

QY 8 FRPDGHLITVAGCGLAELIKGLAYGSDIALDIDOEKTAQAQAEYHKATEELK 67
DB 29 FRLNKGKVASITGSSGIGALAEAFQVADVAIWNSHDAGKALAKKYG----- 82
QY 68 EYPRKGSYACDISDSDYHKKVFAQVAKDFGLPLHLVNTAG--YCENPCECDYPAKNAEK 125
DB 83 ---KVKAYKANVSSSDAVKQTEBOQIKDFGHLIDIVAN-AGIPWTKGAYIDDDDKHPDQ 138
QY 126 MVKNVNLGSLVYS---QAFAPLKEGKGSVVLIGSMGATVNDPQNVVYNNMSKAG 181
DB 139 VVDVDLKGVYAKAGHFRERFEKEGKGA-LVFTASMSGHIVNVQFOATYNAAKAG 197
QY 182 VIHLAKTACEMAKYNIRVNSLNPGYIYGLTKVINGNEELYNRMISGIPQORMSEPK 241
DB 198 VRHFAKSLAVEFAPE-ARVNSVSPGYINTBISDFV---PQETQKNKMSLVPLGRGETAE 253
QY 242 YIGAVLYLSSESASVTTGASLVLDGFT 270
DB 254 LVGAVLFLASD-AGSVATGTDIIVDGGYT 281

RESULT 7

US-09-248-796A-16593
; Sequence 16593, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Ketch Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

```

Query Match      24.3%; Score 344; DB 4; Length 316;
Best Local Similarity 32.0%; Pred No. 2,6e-29;
Matches 87; Conservative 50; Mismatches 107; Indels 28; Gaps 7;

QY      8  PRFDGILITVTGACGGIAGBALIKGLIAYGSDIALDIDDEKTAANKQAEYHKYATBELKUK 67
      |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      66  FSLRDVAVVTAGTGGSIGQTIACSLASAGARVVCFLRDDGGIAETVSH----- 114

QY      68  EYPRKMGSYVC---DISDSDTYHKVVAQAQADPGKPLIHLVNTAGTGCENFPCEEDYPAKKA 123
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      115 -IESIGGQACSYNGVDROAIADLPAVALAKSRGRIDI-AVNAAGIANNPALMESBOW 172

QY      124 EKKNVKNILGSLVLSQAPAKPLIKEGIKASVYLISMSGAIYNDPQONVVYNNMSKAGYI 183
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      173 QRIADITNLGVWNMSCAEALMLKESG--GGSITINISMGGIYNNRGIDQAHNCSAGYI 230

```

```

RESULT 10
US-09-252-991A-30474
; Sequence 30474, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,180
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30474

```

LENGTH: 256
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-30474

Query Match 22.8%; Score 322.5; DB 4; Length 256;
 Best Local Similarity 34.7%; Pred. No. 4,3e-27;
 Matches 93; Conservative 48; Mismatches 104; Indels 23; Gaps 10;

8 FRFDGLTIVTGACGGLAEALIKGLAYGSDIALDIDOEKTAQAQAEYHKYATEELK 67
 8 FDLOEKIAFVGASGIGEBALIKLAQGAHY---VSRKIDGQ-----AVADALTA 58
 QY 68 EVPKKGSYACDISDSTVHKYFAQVAKFGKLPRLVNTAG---YCENFPCEDYPAKXA 123
 DB 59 EGGKATATACHIGEMQIQNFAQIREQFGRLDI-LVNNAATNPQFC-NVLETDLGA--F 114
 QY 124 EKMVYNVLGSLYVQAFAKPLIKGIGASVVLIGSGAIVNDPQOVVYNNKAGVY 183
 DB 115 OKTVDNIRGYFMSIEGSKLMKEHG--GGSIINVASING--VSPGEFGIYSVTKAAYI 170
 QY 184 HLAATLACEMAKYINRVNSLNPGYIYGPITKNVINGNEELVNRWISGIPQGRMSPEKYEY 243
 DB 171 SMTVFPAECAAQFGIRCAALLPGLDTKFAALVK-NDAIRNLALQRIPLKVAEPESEMA 229
 QY 244 GAVLYLSESAASYTTGASLLVDGGFTS 271
 DB 230 GAVLYLASE-ASSYTTGVALNVDGGFLS 256

RESULT 11
 US-09-363-189B-6
 Sequence 6, Application US/09363189B
 Patent No. 6242228
 GENERAL INFORMATION:

APPLICANT: SUGIYAMA, MASAKAZU
 APPLICANT: TONOUCHI, NAOTO
 APPLICANT: SUZUKI, SHUNICHI
 APPLICANT: YOKOZAKI, KENZO
 TITLE OF INVENTION: XYLTOL DEHYDROGENASE OF ACETIC ACID BACTERIA AND GENE THEREOF
 FILE REFERENCE: 0010-1024-0
 CURRENT APPLICATION NUMBER: US/09/363,189B
 CURRENT FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: JP10-216047
 PRIOR FILING DATE: 1998-07-30
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 6
 LENGTH: 262
 TYPE: PRT
 ORGANISM: Gluconobacter oxydans
 US-09-363-189B-6

Query Match 22.7%; Score 321.5; DB 3; Length 262;
 Best Local Similarity 31.0%; Pred. No. 5.8e-27;
 Matches 86; Conservative 53; Mismatches 103; Indels 35; Gaps 10;

9 RFDGLTIVTGACGGLAEALIKGLAYGSDIALDIDOEKTAQAQAEYHKYATEELK 68
 4 KFNKVCIVTGAAGNIGLATLRLAEBSITAIALDMNNEALEKABASYREKGV----- 57
 QY 69 VPKKGSYACDISDSTVHKYFAQVAKFGKLPRLVNTAGYCENF-PCEDYPAKNAEKMV 127
 DB 58 ---ARSYVCVDTSEBAVIGTVDSVVRDGRKIDF-LFNNAAGYQGAFAVQDIPSDDFARVL 113
 QY 128 KVNILGSLYVQAFAKPLIKGIGASVVLIGSGAIVNDPQOVVYNNKAGVYHLAK 187
 DB 114 TINVTGAFAHAKAVSRMITQNY--GRIVNTASWAG--VKPYPMAAAGASGAIIALTE 169
 QY 188 TLACEMAKYINRVNSLNPGYIYGP-----LTKNVINGNE-----ELVNRWISG 231
 DB 170 TVALDLAPYINRVNALSIGV--GGFPMERQVELOAKV--GSQYFSTDIPKVVAAQOMIGSV 226

QY 232 PQGRMSPEKYEYGAVALYLSSESASYYTGASLLVDGG 268
 DB 227 PMRRYDINEIRPGVAFILGDD-SSFMTGVNLPIAGG 262

RESULT 12
 US-09-468-738A-2
 Sequence 2, Application US/09468738A
 Patent No. 6312933
 GENERAL INFORMATION:

APPLICANT: Kimoto, No. 6312933hiro
 APPLICANT: Yamamoto, Hiroaki
 APPLICANT: Mitsunashi, Kazuya
 TITLE OF INVENTION: NOVEL CARBONYL REDUCTASE, METHOD FOR PRODUCING SAID ENZYME, DNA
 TITLE OF INVENTION: ENCODING SAID ENZYME, AND METHOD FOR PRODUCING ALCOHOL USING SAID
 FILE REFERENCE: 06501-050001
 CURRENT APPLICATION NUMBER: US/09/468,738A
 CURRENT FILING DATE: 1999-12-21
 PRIOR APPLICATION NUMBER: JP 1999-171160
 PRIOR FILING DATE: 1999-06-17
 PRIOR APPLICATION NUMBER: JP 1998-363130
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: PatentIn Ver. 2.0, reformatted using WordPerfect 5.1
 SEQ ID NO 2
 LENGTH: 292
 TYPE: PRT
 ORGANISM: Kluyveromyces aestuarii
 US-09-468-738A-2

Query Match 22.1%; Score 312.5; DB 3; Length 292;
 Best Local Similarity 32.3%; Pred. No. 6.9e-26;
 Matches 93; Conservative 50; Mismatches 98; Indels 47; Gaps 10;

3 DYIFTRFDGLTIVTGACGGLAEALIKGLAYGSDIALDIDOEKTAQAQAEYHKYATE 62
 30 NYLSLFGQKGLVITGAGAGIGALCEGFASCGSDVVIID-----YKTSPE 76
 QY 63 ---ELKKEVPMKGSYACDISDSTVHKYFAQVAKFGKLPRLVNT---AGVC----- 110
 DB 77 LGSVLESRYGVRSKYVDITSSSEVDKLVAKILIED---PDRINFVANAGIAMTNGS 133
 QY 111 ---ENFPCEDYPAKNAEKMVYNVLGSLYVQAFAKPLIKGIGASVVLIGSGAIVN 167
 DB 134 IINENATPDVW-----KRWVDVNVQGYTHCAKYAAYFPKQG--HGMILITLASSSTISN 186
 QY 168 DPQOVVYNNKAGVYHLAATLACEMAKY-----NIRVNSLNPGYIYGPITKNVINGNEE 222
 DB 187 VENVQTCYNASKAIVRHMAGFAVEFAHLNTPAKGIRCNISVSPGYTDLTALSAFY---PVR 243
 QY 223 LYNRWISGIPQGRMSPEKYEYGAVALYLSSESASYYTGASLLVDGGFT 270
 DB 244 QRAQMGELTPWGREALPQELVGAITYLASD-AASFNGCDIQVDGGYT 290

RESULT 13
 US-09-940-019-2
 Sequence 2, Application US/09940019
 Patent No. 6416986
 GENERAL INFORMATION:

APPLICANT: Kimoto, No. 6416986hiro
 APPLICANT: Yamamoto, Hiroaki
 APPLICANT: Mitsunashi, Kazuya
 TITLE OF INVENTION: NOVEL CARBONYL REDUCTASE, METHOD FOR PRODUCING SAID ENZYME, DNA
 TITLE OF INVENTION: ENCODING SAID ENZYME, AND METHOD FOR PRODUCING ALCOHOL USING SAID
 FILE REFERENCE: 06501-050001
 CURRENT APPLICATION NUMBER: US/09/940,019
 CURRENT FILING DATE: 2001-08-27
 PRIOR APPLICATION NUMBER: 09/468,738
 PRIOR FILING DATE: 1999-06-17
 PRIOR APPLICATION NUMBER: JP 1998-363130

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:31:08 ; Search time 162 Seconds

(without alignments)
648,653 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415
Sequence: 1 MTDYIPTRFDGHLITVGA.....SAASYTTGASLLVDGPTSW 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubppaa/US10D_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubppaa/US10E_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubppaa/US10F_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubppaa/US11_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 1415 | 100.0 | 272 | US-10-720-018-2 | Sequence 2, Appl1 |
| 2 | 402.5 | 28.4 | 257 | US-10-369-493-2923 | Sequence 2923, Ap |
| 3 | 396 | 28.0 | 211 | US-10-369-493-13036 | Sequence 13036, A |
| 4 | 381 | 26.9 | 283 | US-10-639-159-42 | Sequence 42, Appl |
| 5 | 381 | 26.9 | 283 | US-10-639-159-60 | Sequence 60, Appl |
| 6 | 381 | 26.9 | 283 | US-10-782-258-42 | Sequence 42, Appl |
| 7 | 381 | 26.9 | 283 | US-10-782-258-60 | Sequence 60, Appl |
| 8 | 380 | 26.9 | 283 | US-10-639-159-40 | Sequence 40, Appl |
| 9 | 380 | 26.9 | 283 | US-10-782-258-40 | Sequence 40, Appl |
| 10 | 379 | 26.8 | 283 | US-10-782-258-118 | Sequence 118, Appl |
| 11 | 378 | 26.7 | 283 | US-09-777-157A-1 | Sequence 1, Appl1 |

| | | | | | | |
|----|-------|------|-----|----|---------------------|--------------------|
| 12 | 378 | 26.7 | 283 | 10 | US-09-734-237B-70 | Sequence 70, Appl |
| 13 | 378 | 26.7 | 283 | 16 | US-10-639-159-2 | Sequence 2, Appl1 |
| 14 | 378 | 26.7 | 283 | 16 | US-10-639-159-6 | Sequence 6, Appl1 |
| 15 | 378 | 26.7 | 283 | 16 | US-10-639-159-8 | Sequence 8, Appl1 |
| 16 | 378 | 26.7 | 283 | 16 | US-10-782-258-2 | Sequence 2, Appl1 |
| 17 | 378 | 26.7 | 283 | 16 | US-10-782-258-6 | Sequence 6, Appl1 |
| 18 | 378 | 26.7 | 283 | 16 | US-10-782-258-8 | Sequence 8, Appl1 |
| 19 | 376 | 26.6 | 253 | 15 | US-10-369-493-11548 | Sequence 11548, A |
| 20 | 376 | 26.6 | 283 | 16 | US-10-639-159-46 | Sequence 46, Appl1 |
| 21 | 376 | 26.6 | 283 | 16 | US-10-639-159-54 | Sequence 54, Appl1 |
| 22 | 376 | 26.6 | 283 | 16 | US-10-782-258-4 | Sequence 4, Appl1 |
| 23 | 376 | 26.6 | 283 | 16 | US-10-782-258-46 | Sequence 46, Appl1 |
| 24 | 376 | 26.6 | 283 | 16 | US-10-782-258-54 | Sequence 54, Appl1 |
| 25 | 376 | 26.5 | 283 | 16 | US-10-782-258-116 | Sequence 116, Appl |
| 26 | 375 | 26.5 | 283 | 16 | US-10-782-258-120 | Sequence 120, Appl |
| 27 | 375 | 26.5 | 283 | 16 | US-10-639-159-58 | Sequence 58, Appl1 |
| 28 | 374 | 26.4 | 245 | 15 | US-10-369-493-1304 | Sequence 3304, Ap |
| 29 | 374 | 26.4 | 283 | 16 | US-10-639-159-44 | Sequence 44, Appl1 |
| 30 | 374 | 26.4 | 283 | 16 | US-10-782-258-44 | Sequence 44, Appl1 |
| 31 | 373 | 26.4 | 283 | 16 | US-10-639-159-50 | Sequence 50, Appl1 |
| 32 | 373 | 26.4 | 283 | 16 | US-10-639-159-56 | Sequence 56, Appl1 |
| 33 | 373 | 26.4 | 283 | 16 | US-10-639-159-58 | Sequence 58, Appl1 |
| 34 | 373 | 26.4 | 283 | 16 | US-10-782-258-50 | Sequence 50, Appl1 |
| 35 | 373 | 26.4 | 283 | 16 | US-10-782-258-56 | Sequence 56, Appl1 |
| 36 | 373 | 26.4 | 283 | 16 | US-10-782-258-58 | Sequence 58, Appl1 |
| 37 | 373 | 26.4 | 283 | 16 | US-10-782-258-114 | Sequence 114, Appl |
| 38 | 370 | 26.1 | 283 | 16 | US-10-639-159-52 | Sequence 52, Appl1 |
| 39 | 370 | 26.1 | 283 | 16 | US-10-782-258-52 | Sequence 52, Appl1 |
| 40 | 367 | 25.9 | 251 | 15 | US-10-369-493-14958 | Sequence 14958, A |
| 41 | 367 | 25.9 | 283 | 16 | US-10-639-159-48 | Sequence 48, Appl1 |
| 42 | 367 | 25.9 | 283 | 16 | US-10-782-258-48 | Sequence 48, Appl1 |
| 43 | 364 | 25.7 | 263 | 14 | US-10-128-714-8283 | Sequence 8283, Ap |
| 44 | 362 | 25.6 | 251 | 15 | US-10-369-493-14214 | Sequence 14214, A |
| 45 | 358.5 | 25.3 | 266 | 18 | US-10-508-681-3 | Sequence 3, Appl1 |

ALIGNMENTS

RESULT 1
US-10-720-018-2
; Sequence 2, Application US/10720018
; Publication No. US20040132074A1
; GENERAL INFORMATION:
; APPLICANT: Verbio, Rltva
; APPLICANT: Richard, Peter
; APPLICANT: Penttila, Merja
; TITLE OF INVENTION: New Enzyme for an in vivo and in vitro utilisation of
; TITLE OF INVENTION: Carbohydrates
; FILE REFERENCE: 2530-120
; CURRENT APPLICATION NUMBER: US/10/720,018
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Ambrosiozyma monospora
US-10-720-018-2

Query Match 100.0%; Score 1415; DB 16; Length 272;
Best Local Similarity 100.0%; Pred. No. 5.4e-131;
Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTDYIPTRFDGHLITVGAAGGALIKGLAYGSDIALDDIDQKTAQKQAEYHYKA 60
DB 1 MTDYIPTRFDGHLITVGAAGGALIKGLAYGSDIALDDIDQKTAQKQAEYHYKA 60
QY 1 TEELKLEVPKMGVYACDISDPTVHVPAQVADBEKLPVHTVNTAGYCNFCEYPA 120
DB 61 TEELKLEVPKMGVYACDISDPTVHVPAQVADBEKLPVHTVNTAGYCNFCEYPA 120
QY 121 KNAEKMYVNLGLSLVYQAFAKPLIKGIGKASVVLIGSMGAIIVNDPQNVVYNNNSKA 180

Db 121 KNAERKVVNLGSLVYSGAFAKELIKGASVVLIGSMGSAIVNDPQOVVYNNMSKA 180
Qy 181 GVHIAKLTLACEMAKYINRVNSLNDPGIYVGLTKVINGNEELYNRYISGIPQORMSPK 240
Db 181 GVHIAKLTLACEMAKYINRVNSLNDPGIYVGLTKVINGNEELYNRYISGIPQORMSPK 240
Qy 241 EYIGAVLYLSESAASYTTGASLVDGGFTSM 272
Db 241 EYIGAVLYLSESAASYTTGASLVDGGFTSM 272

RESULT 2
US-10-369-493-2923
; Sequence 2923, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2923
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2923

Query Match 28.4%; Score 402.5; DB 15; Length 257;
Best Local Similarity 36.9%; Pred. No. 5,1e-31;
Matches 101; Conservative 50; Mismatches 90; Indels 33; Gaps 11;
Qy 8 FRPDGHLITVTGACGGALBALIKGLIAYGSDIALDIDDEKTPAKQAEYHKYATEBLK 67
Db 8 FSLKRKVALVTGGGQIGKAIQALAAAGAAVLIMDINE-TARRIVE----- 54
Qy 68 EYFKMS----YACDISDSDYHAKVPAQVAKDPRGLPLVNTAG---YCENPCEPYA 120
Db 55 EIKKGGKADFYVDVYTKEDCEGAVKKALDRWGLDIDG-VNNAIGDWCE--AENYFV 110
Qy 121 KNAERKVVNLGSLVYSGAFAKELIKGASVVLIGSMGSAIVNDPQOVVYNNMSKA 180
Db 111 EKMKKVIDVNLVQ-VFLSKAEFHAKER-KYCKIINIMSMGHIYNNKQKQAVAYASKA 168
Qy 181 GVHIAKLTLACEMAKYINRVNSLNDPGIYVGLTKVINGNEELYNRYISGIPQORMSPK 238
Db 169 GVHILTRSLAAEWAPYRINSGISPIYIRPLIESBNV-----KDLVPLMDLPLGRIGE 224
Qy 239 PKEYIGAVLYLSESAASYTTGASLVDGGFTSM 272
Db 225 VDDLIGAAIFLAS-PASDYVTGHDVLVDGGYTW 257

RESULT 3
US-10-369-493-13036
; Sequence 13036, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2923
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-13036

FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 13036
LENGTH: 211
TYPE: PRT
ORGANISM: Aspergillus nidulans
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(211)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13036

Query Match 28.0%; Score 396; DB 15; Length 211;
Best Local Similarity 41.1%; Pred. No. 1,7e-30;
Matches 88; Conservative 38; Mismatches 64; Indels 24; Gaps 4;

Qy 78 DISDPTVHKVFAQVAKDPGLPLHLVNTAGYCCENPCEDEYPAKNAERKVVNLGSLV 137
Db 1 DVSNDPVSNNALSEVIKAGKID-HLVTSAGFTENFDALISYPRMQKMGVNDGYTLF 59
Qy 138 SQAPAKPLIKGIRGASVVLIGSMGSAIVNDPQOVVYNNMSKGVHIAKLTLACEMAKYN 197
Db 60 ATGVAKHLMERKADG-SIVMIGSMGSAIVNPQOAVNNAKAVRHLASSFAVEMAGHG 118
Qy 198 IRVNSLNDPGIYVGLP-----TKNVINGNEELYNRYISGIPQORM 236
Db 119 IRVNCISPGVWLTALXYPESIQNSXARTICKFVATRRKILIDENPELRDKMISILPGKM 178
Qy 237 SEPKYIGAVLYLSESAASYTTGASLVDGGFT 270
Db 179 GVPELMGAVTFLLSD-ASKYITGADLRVDGYT 211

RESULT 4
US-10-639-159-42
; Sequence 42, Application US/10639159
; Publication No. US20040137585A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Huiteman, Gjalte W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A
; TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
; FILE REFERENCE: 0339.210US
; CURRENT APPLICATION NUMBER: US/10/639,159
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: RaetSEO for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KRED krt215
US-10-639-159-42

Query Match 26.9%; Score 381; DB 16; Length 283;
Best Local Similarity 36.1%; Pred. No. 7,9e-29;
Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8;

Qy 8 FRPDGHLITVTGACGGALBALIKGLIAYGSDIALDIDDEKTPAKQAEYHKYATEBLK 67

```

Db      29  FRLNGKVASITSSSSGIGALAEAPQVADVAITNSQDATGKAEALAKKYG----- 82
      68  EVPRKGSYACDISDSTYHAKFAQVAKDFGLPLHLVNTAG--YCENFPCEBDYPAKNAEK 125
      83  ---KVKAYKANVSSSDAVKQTEIQOIKDFGHLDIIVAN-AGIPMTKGAVIDDQDDDKHFDQ 138
Qy      126  MKVKNLLGSLYVS---QAFAPLKEGIGKASVVLIGSMGATVNDPONOVVNNMSKAG 181
      139  VIDDLKGQVYAKIAGRHRRERPEKEGIGKA-LIFTSVSGHIVNIPOQATYNAAKAG 197
Db      182  VHLAKTLACEMAKYNIRVNSLNPFGYIGPLTKVNINGNEELYNRMISGIPQORMSEPRE 241
      198  VRHFAKSLAVEFAFP-ARVNSVSPGYINTREISDFV---PQETQNMMSLVPLGRGGEFAE 253
Qy      242  YIGAVLYLSEBSAASYTTGASLIVDGGFT 270
      254  LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281
Db

RESULT 5
US-10-639-159-60
; Sequence 60, Application US/10639159
; Publication No. US20040137585A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A.
; TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
; TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
; FILE REFERENCE: 0339.210US
; CURRENT APPLICATION NUMBER: US/10/639,159
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KRED ktr495
US-10-639-159-60

Query Match      26.9%; Score 381; DB 16; Length 283;
Best Local Similarity 36.1%; Pred. No. 7.9e-29;
Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8;
Qy      8  FRPDGHLTIYTGACGGLAEALIKGLAYGSDIALDIDQEKTAQAQAEHYKATBELTK 67
      139  VIIDDLKGQVYAKIAGRHRRERPEKEGIGKA-LIFTSVSGHIVNIPOQATYNAAKAG 197
Db      182  VHLAKTLACEMAKYNIRVNSLNPFGYIGPLTKVNINGNEELYNRMISGIPQORMSEPRE 241
      198  VRHFAKSLAVEFAFP-ARVNSVSPGYINTREISDFV---PQETQNMMSLVPLGRGGEFAE 253
Qy      242  YIGAVLYLSEBSAASYTTGASLIVDGGFT 270
      254  LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281
Db

RESULT 6
US-10-782-258-42
; Sequence 42, Application US/10782258
; Publication No. US20040214297A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A.
; TITLE OF INVENTION: Enzymatic Processes for the Production
; TITLE OF INVENTION: of 4-Substituted 3-Hydroxybutyric Acid Derivatives and
; FILE REFERENCE: 0339.310US
; CURRENT APPLICATION NUMBER: US/10/782,258
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 10/639,159
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KRED ktr215
US-10-782-258-42

Query Match      26.9%; Score 381; DB 16; Length 283;
Best Local Similarity 36.1%; Pred. No. 7.9e-29;
Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8;
Qy      8  FRPDGHLTIYTGACGGLAEALIKGLAYGSDIALDIDQEKTAQAQAEHYKATBELTK 67
      139  VIIDDLKGQVYAKIAGRHRRERPEKEGIGKA-LIFTSVSGHIVNIPOQATYNAAKAG 197
Db      182  VHLAKTLACEMAKYNIRVNSLNPFGYIGPLTKVNINGNEELYNRMISGIPQORMSEPRE 241
      198  VRHFAKSLAVEFAFP-ARVNSVSPGYINTREISDFV---PQETQNMMSLVPLGRGGEFAE 253
Qy      242  YIGAVLYLSEBSAASYTTGASLIVDGGFT 270
      254  LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281
Db

RESULT 7
US-10-782-258-60
; Sequence 60, Application US/10782258
; Publication No. US20040214297A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.

```

```

Db      254  LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281
Db

RESULT 6
US-10-782-258-42
; Sequence 42, Application US/10782258
; Publication No. US20040214297A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A.
; TITLE OF INVENTION: Enzymatic Processes for the Production
; TITLE OF INVENTION: of 4-Substituted 3-Hydroxybutyric Acid Derivatives and
; FILE REFERENCE: 0339.310US
; CURRENT APPLICATION NUMBER: US/10/782,258
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 10/639,159
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KRED ktr215
US-10-782-258-42

Query Match      26.9%; Score 381; DB 16; Length 283;
Best Local Similarity 36.1%; Pred. No. 7.9e-29;
Matches 97; Conservative 52; Mismatches 96; Indels 22; Gaps 8;
Qy      8  FRPDGHLTIYTGACGGLAEALIKGLAYGSDIALDIDQEKTAQAQAEHYKATBELTK 67
      139  VIIDDLKGQVYAKIAGRHRRERPEKEGIGKA-LIFTSVSGHIVNIPOQATYNAAKAG 197
Db      182  VHLAKTLACEMAKYNIRVNSLNPFGYIGPLTKVNINGNEELYNRMISGIPQORMSEPRE 241
      198  VRHFAKSLAVEFAFP-ARVNSVSPGYINTREISDFV---PQETQNMMSLVPLGRGGEFAE 253
Qy      242  YIGAVLYLSEBSAASYTTGASLIVDGGFT 270
      254  LVGAYLFLASD-AGSYATGTDTIIVDGGYT 281
Db

RESULT 7
US-10-782-258-60
; Sequence 60, Application US/10782258
; Publication No. US20040214297A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Hulsman, Gjalte W.
; APPLICANT: Ma, Steven K.

```

```

APPLICANT: Newnan, Lisa M.
APPLICANT: Sheldon, Roger
APPLICANT: Wang, Li A.
TITLE OF INVENTION: Enzymatic Processes for the Production
TITLE OF INVENTION: of 4-Substituted 3-Hydroxybutyric Acid Derivatives and
TITLE OF INVENTION: Vicinal Cyano, Hydroxy-substituted Carboxylic Acid Ester
FILE REFERENCE: 0339, 310US
CURRENT APPLICATION NUMBER: US/10/782,258
CURRENT FILING DATE: 2004-02-18
PRIOR APPLICATION NUMBER: US 10/639,159
PRIOR FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: US 60/402,436
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 126
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60
LENGTH: 283
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: KRED krh495
US-10-782-258-60

Query Match          26.9%; Score 381; DB 16; Length 283;
Beet Local Similarity 36.1%; Pred. No. 7,9e-29;
Matches 97; Conservative 52; Mismatches 98; Indels 22; Gaps 8

QY      8 FRFDGHLITVTCAGCGLAELIKGLIAYGSDIALDIDDEKTRAKQAEYHKATETSLK 67
DB      29 FPLNGKVASITSSSGIGYALAEFAQVAGADVAIWNSDADGKKEALAKKGV----- 82
QY      68 EYPRKGSYACDISDSDPTHKVFAQYAKKDGKGLPLHLVNTAG--YCFNFCEDVPAPAKNR 125
DB      83 ---KRYAYANVSSSDPAVQOTTEBOQIKDFGHLDIVAN--AGIPWYGAITYDDDDKHFPQ 138
QY      126 WYKNVLGSLYS---QAFAKPLIEKIGKASVVLIGMSGAIYNDPQNVVYNNKSG 181
DB      139 VIDVDLKGVGYAKTAGKRRYRERFEKIGKA-LITFASVSGHIVNIPQGAITYNAKKG 197
QY      182 VIHLAKTLCEWAKNIRVNSLNPGYTYGPTKVNINGNEELYNNRISGIPQORMSPKE 241
DB      198 VAFPAKSLAVEFAPF--ARVNSVSPGYINTEISDFV---PQETONKWSLVPLGRGERTAE 253
QY      242 YTGAVLYLSEGAASYTTGASLLVDGFT 270
DB      254 LVGAYVLEFLASD--AGSYATGTHDIVDGXYT 281

RESULT 8
US-10-639-159-40
/ Sequence 40, Application US/10639159
/ Publication No. US20040137585A1
GENERAL INFORMATION:
APPLICANT: Davis, S. Christopher
APPLICANT: Grate, John H.
APPLICANT: Gray, David R.
APPLICANT: Gruber, John M.
APPLICANT: Huisman, Gjal't W.
APPLICANT: Ma, Steven K.
APPLICANT: Newman, Lisa M.
APPLICANT: Sheldon, Roger
APPLICANT: Wang, Li A
TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
FILE REFERENCE: 0339, 210US
CURRENT APPLICATION NUMBER: US/10/639,159
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: US 60/402,436
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 40
LENGTH: 283

```

```

; TYPE: PRT ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: KRED krh133c US-10-639-159-40
Query Match          26.9%; Score 380; DB 16; Length 283;
Best Local Similarity 36.8%; Pred. No. 9,9e-29;
Matches      99; Conservative 49; Mismatches 99; Indels 22; Gaps 8;

QY    8 FRPDGHLTIYTGCAGCGLAELIKGLAYGSIDALLDDIOEKTAKKQAEEHYHATATEBLTKL 67
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB    29 FKLNGKVASITSGSSGGICGYALAEFAOVGVADVAILWNSODATGAELAAXTYV----- 82
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY    68 EVPKMGSACDISDSDTHKHVPAOVAKFGLPLHLVNTAG--YCENFPCEDPYAKNAEK 125
       ::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB    83 --KVAKKAVNWSDDDAVKQTIEOIKDFGHLDIVAN-AIGPMYKGAVIQDDDKHPDQ 138
       :||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY    126 WKVNLLGSLYVS---QAFAPLKKEGIKQASVVLLIGSMGAIYNPQPONVYVMSTAQ 181
       :||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB    139 VVDVDVLKVGYYAAKHAIGHRYERFEKEGKKA-LVFATSMGHIYNNVPOFQATYMAAKG 197
       :||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY    162 VIHIAKTACEWAKTNIRVNSLNPFITYGLPTKNVINGNNEELYNMWISIGIPQORMSPEPK 241
       :||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB    198 VHFFAKSLAVEAAPP-ARVNSVSPCYINTSIDFV---PQETQNKMSSLVPLGRGETRAE 253
       :||::||::||::||::||::||::||::||::||::||::||::||::||::||
QY    242 YTGAVLYLSBSASAFTTGTASLLVNDGGFT 270
       :||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB    254 LVGAVLYFLASD-AGSYIATGTDIIVDGFT 281
       :||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 9
US-10-782-258-40
; Sequence 40, Application US/10782258
; Publication No. US20040214297A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.
; APPLICANT: Gruber, John M.
; APPLICANT: Huileman, Gjalt W.
; APPLICANT: Ma, Steven K.
; APPLICANT: Newman, Lisa M.
; APPLICANT: Sheldon, Roger
; APPLICANT: Wang, Li A.
; TITLE OF INVENTION: Enzymatic Processes for the Production
; TITLE OF INVENTION: of 4-Substituted 3-Hydroxybutyric Acid Derivatives and
; TITLE OF INVENTION: Vicinal Cyano, Hydroxy-substituted Carboxylic Acid Esters
; FILE REFERENCE: 0339.310US
; CURRENT APPLICATION NUMBER: US/10/782,258
; CURRENT FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: US 10/639,159
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 60/402,436
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: KRED krh133c
US-10-782-258-40

Query Match          26.9%; Score 380; DB 16; Length 283;
Best Local Similarity 36.8%; Pred. No. 9,9e-29;
Matches      99; Conservative 49; Mismatches 99; Indels 22; Gaps 8;

QY    8 FRPDGHLTIYTGCAGCGLAELIKGLAYGSIDALLDDIOEKTAKKQAEEHYHATATEBLTKL 67
       :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB    29 FKLNGKVASITSGSSGGICGYALAEFAOVGVADVAILWNSODATGAELAAXTYV----- 82

```

```

QY 68 EYFKKGSYACDSDSPTTHKVFNOVAKDGGKLEPHLVN7AG--YCNENFPEEDYPAKNAEK 125
Db 83 ---KKRAYANWSSSDAVAKQITEBOQIKDKEGHDIVAN-AGLEPW7GAYTDQDDDKHFQD 138
QY 126 MYKNVNLGSLYYVS---QAFAPPLIKEG1KGASVVLIGSMSGAIIVNDPONOVVYNNMSKAG 181
Db 139 YVDVDLKGGVYAKAGRHRYRERFEKEGKKGA-LVFTASMSGH1VAVPQGAATYNNAAKAG 197
QY 182 VTHLAKTACEAKNIRVNSLNPGYTYGPTKYNVINGNEELYNRMISG1PQORMSEPKX 241
Db 198 VWHFPAKSLAVEBAPF-ARVNSVSPG1YINNEISDFV---PQETONKWSLVLPLRGGETAE 253
QY 242 Y1GAVLV1SEBASAYTTGASLLVDDGGFT 270
Db 254 LVGAYLVFLASD-AGSYATGTDT1IVDDGYT 281

```

```

RESULT 10
US-10-782-258-118
Sequence 118, Application US/10782258
Publication No. US20040214297A1
GENERAL INFORMATION:
APPLICANT: Davis, S. Christopher
APPLICANT: Grace, John H.
APPLICANT: Gray, David R.
APPLICANT: Gruber, John M.
APPLICANT: Hulsman, Gjalte W.
APPLICANT: Ma, Steven K.
APPLICANT: Newman, Lisa M.
APPLICANT: Sheldon, Roger
APPLICANT: Wang, Li A.
TITLE OF INVENTION: Enzymatic Processes for the Production
of 4-Substituted 3-Hydroxybutyric Acid Derivatives and
TITLE OF INVENTION: Vicinal Cyano, Hydroxy-substituted Carboxylic Acid Esters
FILE REFERENCE: 0359.310US
CURRENT APPLICATION NUMBER: US/10/782,258
CURRENT FILING DATE: 2004-02-18
PRIOR APPLICATION NUMBER: US 10/639,159
PRIOR FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: US 60/402,436
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 126
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 118
LENGTH: 283
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: KRED S01091625
US-10-782-258-118

```

| | | | | |
|-----------------------|------------------|-------------------|------------|-------------|
| Query Match | 26.8% | Score 379; | DB 16; | Length 283; |
| Best Local Similarity | 37.2% | Pred No. 1.2e-28; | | |
| Matches 100; | Conservative 47; | Mismatches 100; | Indels 22; | Gaps 8 |

Qy 8 FRDGHVITVIGACGGIAALIKGLLAAYSIDIALDIDQEKYAAQAVHYRYATEELK 67
Db 29 FKJNGVVASITGNSGIGYALAEAFQAVADVAIWNNSHDATGKAEALAKKYGV----- 82
Qy 68 EYFKMSYACDIDSPYHKVFQAQAKDEGKPLHJVNTAG--YOENPCEDYAPAKNAEK 12
Db 83 ---KVAAYKANYSSSNAVKQITIEQIKDQPGHLDIVVAN-AGIPTWKAYITDDDKDKFHQ 13
Qy 126 MYKYNLLGSLIYS---QAPAKPLIKEGIKGASVVLIGSMGSAIVNDPQNVYNYMSKAG 181
Db 139 VVDVDLKGAGYAKAHGRRFRFRFEKEGKGA-LVFTASMSGHIIWVYQFPATYNNAVAG 197
Qy 182 VIHAKTLTCEWAKYINRVNSLNPGYIPLTKYINNGEELVYRWISIGIPQOMSEPK 241
Db 198 VRFASLAVEPFP-ARVNSVSPGICINIEIDFV---PQETQRNWSLVFLRGGETAE 253
Qy 242 YIGAVIYLLSESAASTTGAALLVDGGFT 270

Db 254 LVGAYLFLASD-AGSYATGTDIIVDGGYT 281

```

RESULT 11
US-09-777-157A-1
; Sequence 1, Application US/09777157A
; Publication No. US20020006651A1
; GENERAL INFORMATION:
; APPLICANT: Yasohara, Yoshihiko
; APPLICANT: Kizaki, No. 6448052jyuk1
; APPLICANT: Hasegawa, Junzo
; APPLICANT: Wada, Masaru
; APPLICANT: Shimizu, Sakayu
; APPLICANT: Katoka, Michihiko
; APPLICANT: Yamamoto, Kazuhiko
; APPLICANT: Kawabata, Hiroshi
; APPLICANT: Kita, Keiko
; TITLE OF INVENTION: Carbonyl Reductase Enzyme and Methods for its Use
; FILE REFERENCE: 068383, 0110
; CURRENT APPLICATION NUMBER: US/09/777,157A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 09/367,012
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Candida magnoliae
; US-09-777-157A-1

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 26.7%; | Score 378; | DB 9; | Length 283; |
| Best Local Similarity | 36.8%; | Pred. No. 1.6e-28; | | |
| Matches 99; | Conservative 48; | Mismatches 100; | Indels 22; | Gaps 8; |

```

RESULT 12
US-09-734-237B-70
; Sequence 70, Application US/09734237B
; Publication No. US20030064432A1
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 70

```

```

; LENGTH: 283
; TYPE: PRT
; ORGANISM: Candida magnoliae
US-09-734-237B-70

```

| | | | | |
|-----------------------|--------|--------------------|-----------------|-------------|
| Query Match | 26.7%; | Score 378; | DB 10; | Length 283; |
| Best Local Similarity | 36.8%; | Pred. No. 1.6e-28; | | |
| Matches | 99; | Conservative 48; | Mismatches 100; | Indels 22; |
| | | | | Gaps 8; |

QY 8 PRFDHULTVTAACGGIALAETKGLLAYSDDLIDIDDEKTPAKQAEYHKYATEELKUK 67
 Db 29 FCLNGKVASITSSSSGIGYALAEAPQVADVAIWNSHDAIKABALAKKTV----- 82
 QY 68 EYFKMGSAYCIDSIDPTVHKVFAQVAKDPGLPLHLVNTAG--YCENFPCEDEYPAKNAEK 125
 Db 83 ---KWAYANANSSSPAVNKOTIEOQIKDGHIDIYVAN-AGIPWTKGAYIIDODDDHFFQ 138
 QY 126 MYKNVLLGSLUYS---QAFAPKPLIEGFGKASVVLIGSMGAIYVNDPQNVUYNNSKAG 181
 Db 139 VDVVDIKGVGYAKAGAGHFRFRFEKGGKGA-LVYTAASMSGHIVVPOQATYNAAKKG 197
 QY 182 VTHLAKTACEWAKNIRVNSLNPGYIYGPILTNTVINGNEELYNRMISGIPQORMSEPKYE 241
 Db 198 VAKPAKSLAVEPAPF-ARVNSVSPGYINTEISDFV---PQETONKWSLTVPLRGGEITAE 253
 QY 242 YTGAVLVYLLSESAASYTTGASLLYDDGFT 270
 Db 254 LVGAYLFLASD-AGSYATGCTDIIIVDDGGYT 281

RESULT 13
US-10-639-159-2

| | Query Match | 26.7%; | Score 378; | DB 16; | Length 283; |
|----|-----------------------|--|-------------------|-----------------|-------------------|
| | Best Local Similarity | 36.8%; | Pred. No.1.6e-28; | | |
| | Matches | 99; | Conservative 48; | Mismatches 100; | Indels 22; Gaps 8 |
| QY | 8 | FREDHLITVTAQCGSLAEALIKGLIAYSDIALDIDDEKTAQAQAEYHKVATSELKX | 67 | | |
| DB | 29 | FKLNGKVASITGSSSGIGYALAEFAQVADVAIWNSHDAQKAEALAKKTV----- | 82 | | |
| QY | 68 | EYFKMSYACDISDSTYHKVFAQVAKDEGKLEPLHLVNTAG--YCENFPCEQDYPAPAKMAEK | 125 | | |
| DB | 83 | ---KVAKYANANSSSDPAVKQTIEQIKDFGHLDIIVAN--AGIPWTGAYIIDDDDKGHFQ | 138 | | |
| QY | 126 | MYKVNLLGSLIYS---QAFAKPLIEEGIKGASVVLIGSMGALIVNDPONOVVYNNMSKGG | 181 | | |

Db 139 VDVVDLKGVGYYAKNAHGRRFRFRFEKEKKKA-LVFATSMGCHIVNPOFOATYNAAKG 197

QY 182 YIHATLACENAKXNIVNLSNIPGYITGFLKXVINGNEELVYRMTSGIPQORMSPKE 241

Db 198 VRHFAASLAEFAPF-ARVNVSFGYINTEISDF---PQLTKKMWSLVPLRGGEYAE 253

QY 242 YIGAVLYLISESAASYTTGASLLVNGGTF 270

Db 254 LVGAYVLFPLASD-AGSYATGTDTILLVDGGYI 281

```

RESULT 14
US-10-639-159-6
/ Sequence 6, Application US/10639159
/ Publication No. US20040137585A1
/ GENERAL INFORMATION:
/ APPLICANT: Davis, S. Christopher
/ APPLICANT: Grate, John H.
/ APPLICANT: Gray, David R.
/ APPLICANT: Gruber, John M.
/ APPLICANT: Hulsman, G'alt W.
/ APPLICANT: Ma, Steven K.
/ APPLICANT: Newman, Lisa M.
/ APPLICANT: Sheldon, Roger
/ APPLICANT: Wang, Li A
/ TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
/ TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
/ FILE REFERENCE: 0339.210US
/ CURRENT APPLICATION NUMBER: US/10/639,159
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: US 60/402,436
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 92
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 283
/ TYPE: prt
/ ORGANISM: Candida magnoliae
US-10-639-159-6

```

| Query Match | 26.7% | Score 378 | DB 16 | Length 283 |
|-------------|-------|--|-------|----------------|
| Similarity | 36.8% | Pred. No. 1.6e-28 | | |
| Matches | 99 | Conservative | 48 | Mismatches 100 |
| | | | | Indels 22 |
| | | | | Gaps 8 |
| QY | 8 | PPEDDHLITVTCAGCGLAALIKGILAAAGSDIALDIDIOEKTPAAKQAEVHKYATBELXK | 67 | |
| DB | 29 | FTLNGKVASITSSSSGIGYALLAAEAQUGADVALIYNSHDATGKAEALAKKYG----- | 82 | |
| QY | 68 | EYPMKMSYACDIDSDDTYHKPNOYAKDFGKLPHLVNTAG--YCENFPCEDYPAKNAEK | 125 | |
| DB | 83 | ---KYKAYANVSSSDAVAKQTIEQDIKDFGHLDIVAN-AGIPMTKGAYITDQDDDKHPDQ | 138 | |
| QY | 126 | MVKNNILGSLYVS---QAFAKPLIKEGIKGASVVLISMSGALVYNDQNOVYVYMSKRG | 181 | |
| DB | 139 | VVDVVLKGVYAKHAKGAFHFRERFEKEGKKGA-LVFTASMSGHIVNVPOFQATVYAAARG | 197 | |
| QY | 182 | VYHLAKTLACBFAKKNIRVNSLNPETIYGPFLKTVNINSGEELVNRMISGIPQORMSEPR | 241 | |
| DB | 198 | VHFPAKSLAVEAPF-ARVNSVSPGIVNTEISDFV--PQETQNMKMSLVPLRGGEITAE | 253 | |
| QY | 242 | YTGAVLYLISESAASYTTGASLLVDDGGFT | 270 | |
| DB | 254 | LWGAVYFLASD-AGSYATGTDLITVGGYT | 281 | |

RESULT 15
US-10-639-159-8
; Sequence 8, Application US/10639159
; Publication No. US2004013785A1
; GENERAL INFORMATION:
; APPLICANT: Davis, S. Christopher
; APPLICANT: Grate, John H.
; APPLICANT: Gray, David R.

```

1  APPLICANT: Gruber, John M.
2  APPLICANT: Hulsmann, Gjaalt W.
3  APPLICANT: Ma, Steven K.
4  APPLICANT: Newman, Lisa M.
5  APPLICANT: Sheldon, Roger
6  APPLICANT: Wang, Li A
7  TITLE OF INVENTION: ENZYMATIC PROCESSES FOR THE PRODUCTION
8  OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID DERIVATIVES
9  TITLE OF INVENTION: OF 4-SUBSTITUTED-3-HYDROXYBUTYRIC ACID
10 FILE REFERENCE: 0339,210US
11 CURRENT APPLICATION NUMBER: US/10/639,159
12 CURRENT FILING DATE: 2003-08-11
13 PRIOR APPLICATION NUMBER: US 60/402,436
14 PRIOR FILING DATE: 2002-08-09
15 NUMBER OF SEQ ID NOS: 92
16 SOFTWARE: FastSeq for Windows Version 4.0
17 SEQ ID NO 8
18 LENGTH: 283
19 TYPE: PRT
20 ORGANISM: Candida magnoliae
21 US-10-639-159-8

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query March | 26.7% | Score 378; | DB 16; | Length 283; |
| Best Local Similarity | 36.8%; | Pred. No. 1.6e-28; | | |
| Matches 99; | Conservative 48; | Mismatches 100; | Indels 22; | Gaps 8 |

Search completed: July 9, 2005, 10:44:43
Job time : 163 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:20:26 ; Search time 41 Seconds

(without alignments)
638.316 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415

Sequence: 1 MTDIPTFRFGHLITVTGA.....SAASYTGASLLVDGGFTSW 272

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 798.5 | 56.4 | 278 | 2 | D-arabinitol 2-deh |
| 2 | 797 | 56.3 | 282 | 2 | D-arabinitol 2-deh |
| 3 | 402.5 | 28.4 | 257 | 2 | oxidoreductase, sh |
| 4 | 378 | 26.7 | 283 | 2 | cardonyl reductase |
| 5 | 376 | 26.6 | 262 | 2 | short chain dehydr |
| 6 | 376 | 26.6 | 262 | 2 | oxidoreductase, sh |
| 7 | 361 | 25.5 | 255 | 2 | borbicol utilizati |
| 8 | 358.5 | 25.3 | 281 | 2 | 2-deoxy-D-gluconat |
| 9 | 349 | 24.7 | 271 | 2 | AC0157 |
| 10 | 344 | 24.3 | 255 | 2 | D70635 |
| 11 | 336 | 23.7 | 255 | 2 | oxidoreductase, sh |
| 12 | 332.5 | 22.8 | 255 | 2 | D83416 |
| 13 | 331.5 | 22.0 | 262 | 2 | JC7939 |
| 14 | 312 | 22.0 | 255 | 2 | TJ5987 |
| 15 | 305.5 | 21.6 | 267 | 2 | AH2352 |
| 16 | 305 | 21.6 | 261 | 2 | TJ8157 |
| 17 | 303.5 | 21.4 | 271 | 2 | D84695 |
| 18 | 302 | 21.3 | 254 | 2 | A11117 |
| 19 | 302 | 21.3 | 254 | 2 | AC1478 |
| 20 | 301 | 21.3 | 268 | 2 | C84693 |
| 21 | 298 | 21.1 | 251 | 2 | P95295 |
| 22 | 291 | 20.6 | 263 | 2 | AB0812 |
| 23 | 289 | 20.4 | 256 | 2 | A57149 |
| 24 | 289 | 20.4 | 521 | 1 | C87474 |
| 25 | 288.5 | 20.3 | 260 | 2 | C87408 |
| 26 | 287.5 | 20.3 | 262 | 2 | B84695 |
| 27 | 287.5 | 20.3 | 268 | 2 | B84695 |
| 28 | 287 | 20.3 | 298 | 2 | A13058 |
| 29 | 287 | 20.3 | 298 | 2 | G98227 |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------------|
| 30 | 285.5 | 20.2 | 262 | 2 | C84711 | probable tropinone |
| 31 | 285 | 20.1 | 256 | 2 | A10406 | probable dehydroge |
| 32 | 284 | 20.1 | 255 | 2 | G82644 | 2,5-dichloro-2,5-c |
| 33 | 281.5 | 19.9 | 262 | 2 | C84694 | probable tropinone |
| 34 | 281.5 | 19.9 | 322 | 2 | D84694 | probable tropinone |
| 35 | 281 | 19.9 | 246 | 2 | H72219 | 3-oxoacyl-(acyl ca |
| 36 | 281 | 19.9 | 268 | 2 | G96016 | probable gluconate |
| 37 | 279 | 19.7 | 285 | 1 | A65017 | probable 3-oxoacyl |
| 38 | 278 | 19.6 | 269 | 2 | C84695 | probable tropinone |
| 39 | 277 | 19.6 | 268 | 2 | B84693 | probable tropinone |
| 40 | 277 | 19.6 | 285 | 2 | A91041 | probable oxidoredu |
| 41 | 277 | 19.6 | 285 | 2 | D85885 | probable oxidoredu |
| 42 | 276.5 | 19.5 | 253 | 2 | C91091 | 2-deoxy-D-gluconat |
| 43 | 275.5 | 19.5 | 253 | 2 | F85936 | 2-deoxy-D-gluconat |
| 44 | 275 | 19.4 | 267 | 2 | G97220 | short-chain alcoh |
| 45 | 274.5 | 19.4 | 260 | 2 | A84695 | probable tropinone |

ALIGNMENTS

RESULT 1

S57351 D-arabinitol 2-dehydrogenase (EC 1.1.1.-) - Yeast (Pichia stipitis)

C:Species: Pichia stipitis

C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004

C:Accession: S57351; S57352

R:Hallborn, J.; Walfriedson, M.; Penttilae, M.; Keranen, S.; Hann-Haegrdal, B.

A>Title: A short-chain dehydrogenase gene from Pichia stipitis having D-arabinitol dehydr

A:Reference number: S57351; MUID:96090133; PMID:7483848

A:Accession: S57351

A:Molecule type: DNA

A:Residues: 1-278 <HML>

A:Cross-references: UNIPROT:P50167; EMBL:Z46866; NID:G758241; PIDN:CAA86939.1; PID:G7631

A:Accession: S57352

A:Molecule type: mRNA

A:Residues: 1-278 <HAM>

A:Cross-references: EMBL:Z46866; NID:G758241; PIDN:CAA86939.1; PID:G763164

A:Experimental source: strain CBS 6054

C:Genetics:

A:Gene: ARDH

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

C:Keywords: NAD; oxidoreductase

F:17-212/Domain: short-chain alcohol dehydrogenase homology <SADH>

F:17-47/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 56.4%; Score 798.5; DB 2; Length 278;
Best Local Similarity 54.3%; Pred. No. 1.1e-55;
Matches 151; Conservative 54; Mismatches 64; Indels 9; Gaps 3;

| | | | |
|----|-----|--|-----|
| QY | 3 | DY-----IPPRFGCHLTIVTGACGGLAEALIKLLAVGSLALDDIDDEKTAQKQAEYH | 57 |
| DB | 2 | DYSYANVVPNFRLLDGRLLAIIITGSGGLAAVLSRALLOGADVALLDNLEETKKAANEVL | 61 |
| QY | 58 | KVATESELKKEVPMG---SYACDISPSDPTVHKVFAOVAQDFGLPPLVVTAGCENFP | 114 |
| DB | 62 | GWGSEITLKGHAAISIGVVSAMSCNIGDAEAVDAFFSSINEHGKIDALLITTAGICENFP | 121 |
| QY | 115 | CEDYPANNAEKVKNLLGLSYQAFAKPLIKEGIGASVVLIGSNGAIVNDPQNV | 174 |
| DB | 122 | AETYPATNAESIMKVMGIGSFYVSQSPARPLIQNNLNG-SIILIGMSGITVNDPQNCM | 180 |
| QY | 175 | YNNKAGVIHLAKTLACEMAKVNRUNSLMPGYTGGTLKRVNNGNBNELNRYWISGIFQ | 234 |
| DB | 181 | YNNKAGVIHLVRSLLACEMAKVNRVNTLSPGYTLPLTRNVISGHTENKAEWSKIPMK | 240 |
| QY | 235 | RMSEKVEYIGAVLYLSESAASYTTGASLLVDGGFTSW | 272 |
| DB | 241 | RMSEKVEYISLYLASETASSTYTGNNLVVDGGYECW | 278 |

RESULT 2

RESULT 5
AG2809
Short chain dehydrogenase Atu1897 [Imported] - Agrobacterium tumefaciens (strain C58, Du
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AG2809
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kucyavin, T.; Levy, R.; Li, M.; Mclell
; Karp, P.; Romero, P.; Zhang, S.
Science 294 2317-2323, 2001
A:Authors: Yoo, H.; Tso, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
eter, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG2809
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <KUR>
A:Cross-references: UNIPROT:Q8UB64; GB:AE008688; PIDN:AAL42893.1; PID:g17740346; GSPDB:C
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu1897
A:Map position: circular chromosome
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.6%; Score 376; DB 2; Length 262;
Best Local Similarity 35.8%; Pred. No. 2.5e-22;
Matches 96; Conservative 43; Mismatches 109; Indels 20; Gaps 7;

QY 8 FRPDGHLTIYVAGCGGLAELIKGLAYGSDIALDIDOEKTAQAQAYHKAATBELKX 67
DB 12 FDLGQVALVTGAGSGIGQRIAMGLAQGANVALLDRRTDGLAQTAQFIARAG----- 66
QY 68 EVPKMGYACISDSDTYHKAFAQVAKDFGLPLHLVNTAGYCNFPCEDYPAKAEKV 127
DB 67 ---KSIQIAADVTSKQALTDVAARQAELGLSL-AVNAAGIANNPABEEMESQFQTM 122
QY 128 KVNLLGSLYVQAFAKPLIKGKIGKASVLLIGSMGALVNDPQNVVYMSKAGVYHLAK 187
DB 123 DINKGVFLSCQAEANMLKNG-RGA-IVNIAISMGVIVNGLMQCHYNASAGVYHMSK 180
QY 188 TLACEMAKYINRVNSLNGYIYGPLTKVINGNEELVNR---WISGIPQRMSEPKYIG 244
DB 181 SMAMWVGRIKRVNTISPGYATP-----NTRPEMVHQTFLFEQTMQMAGVDENVG 235
QY 245 AVLYLLSESAASYTTGASLLVDGFTSW 272
DB 236 PAIFLLSD-AASFVTVGVLDLVGGFCCW 262

RESULT 6
B97588
oxidoreductase, short chain dehydrogenase/reductase family [Imported] - Agrobacterium tu
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97588
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B97588
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <KUR>
A:Cross-references: UNIPROT:Q8UB64; GB:AE007869; PIDN:AAK97659.1; PID:g15157014; GSPDB:C
C:Genetics:
A:Gene: AGR C 3482
A:Map position: circular chromosome
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.6%; Score 376; DB 2; Length 262;
Best Local Similarity 35.8%; Pred. No. 2.5e-22;
Matches 96; Conservative 43; Mismatches 109; Indels 20; Gaps 7;

QY 8 FRPDGHLTIYVAGCGGLAELIKGLAYGSDIALDIDOEKTAQAQAYHKAATBELKX 67
DB 12 FDLGQVALVTGAGSGIGQRIAMGLAQGANVALLDRRTDGLAQTAQFIARAG----- 66
QY 68 EVPKMGYACISDSDTYHKAFAQVAKDFGLPLHLVNTAGYCNFPCEDYPAKAEKV 127
DB 67 ---KSIQIAADVTSKQALTDVAARQAELGLSL-AVNAAGIANNPABEEMESQFQTM 122
QY 128 KVNLLGSLYVQAFAKPLIKGKIGKASVLLIGSMGALVNDPQNVVYMSKAGVYHLAK 187
DB 123 DINKGVFLSCQAEANMLKNG-RGA-IVNIAISMGVIVNGLMQCHYNASAGVYHMSK 180
QY 188 TLACEMAKYINRVNSLNGYIYGPLTKVINGNEELVNR---WISGIPQRMSEPKYIG 244
DB 181 SMAMWVGRIKRVNTISPGYATP-----NTRPEMVHQTFLFEQTMQMAGVDENVG 235
QY 245 AVLYLLSESAASYTTGASLLVDGFTSW 272
DB 236 PAIFLLSD-AASFVTVGVLDLVGGFCCW 262

RESULT 7
T39164
sorbitol utilization protein soul - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39164
R:McLean, J.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21831
A:Accession: T39164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-255 <MCL>
A:Cross-references: UNIPROT:Q9Y629; EMBL:AL021817; PIDN:CAB40197.1; GSPDB:GN00066; SPDB:
A:Experimental source: strain 972h-, cosmid c8E11
C:Genetics:
A:Gene: SPDB:SPAC8E11.10
A:Map position: 1
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 25.5%; Score 361; DB 2; Length 255;
Best Local Similarity 36.6%; Pred. No. 3.7e-21;
Matches 98; Conservative 45; Mismatches 103; Indels 22; Gaps 10;

QY 8 FRPDGHLTIYVAGCGGLAELIKGLAYGSDIALDIDOEKTAQAQAYHKAATBELKX 67
DB 5 FSLKGTTLITGSGGIGFSTAKAPAAAGSNVGL-YGRNKKALFYA-----AELRDK 56
QY 68 EVPKMGYACISDSDTYHKAFAQVAKDF-GKPLHLVNTAGYCNFPCEDYPAKAEKV 124
DB 57 HGVQAKAYSCPIERSAVIETTNQAVBELGRLDVIAMNGIALPHLSLD---KNEDIM 113
QY 125 -KMKVNLGSLYVQAFAKPLIKGKIGKASVLLIGSMGALVNDPQNVVYMSKAGVY 183
DB 114 TKVGVINLNGYVYAAQAGHFKQKQ-KG-SLFTFASMGHIANWPQOMASYHATKAVK 171
QY 184 HLAATLCEMAKYINRVNSLNGYIYGPLTKVINGNEELVNRWISGIPQRMSEPKYI 243
DB 172 HLAALAVENAPF-ARNSVSPGIITDITL---LYADENLRKKKKEYTPQARIGLPDEL 227
QY 244 GAVLYLLSESAASYTTGASLLVDGFTS 271
DB 228 GAVLYLLSD-AASYCTGSDIIVDGYCS 254

RESULT 8
F69400
2-deoxy-D-gluconate 3-dehydrogenase (kduD) homolog - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus
 C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C/Accession: F69400
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Overbeek, R.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: F69400
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-281 <KLE>
 A:Cross-references: UNIPROT:O29061; GB:AEO01021; GB:AEO00782; NID:g2689344; PIDN:AAB9004
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 F/36-216/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 25.3%; Score 358.5; DB 2; Length 281;
 Best Local Similarity 33.3%; Pred. No. 6-6e-21;
 Matches 91; Conservative 53; Mismatches 106; Indels 23; Gaps 9;
 QY 4 YIPFRDGLTIVTGACGGLAELIKGLAYGSDIALDIDQEKTAQAQEHKATAE 63
 DB 27 YRKMFDLGTVAIVTGATGCGPGLADPGCDVVVV-----GRRLD-----VLEK 74
 QY 64 LKAKVPMGSA-----CDISDSTVHVFPVQVAKDFGKLLHLVNTGCENPFCEY 119
 DB 75 LK-SIEKLGORALAVKCDITSEEDVANLVNRYVEEFGRIDI-LVNCGINIPKPAEY 132
 QY 120 AKNAEKVKNVLGSLVYSQAFKPLKEGIGASVVLIGMSGAIIVDPQNVVYNSK 179
 DB 133 LEDNKKVNDVNTGVFLVCREVGKRMVQ--NGGKIIVSSVRS--YGMKNYLAAGCSK 189
 QY 180 AGVHLAKTLACEMAKNIRVNSLNPXYIGPLTKVINGBELYNNISGIPQGRSE 239
 DB 190 AAVNMITQLACEMAKNIVLVAIPVIAPIPLTHIMK-DPELSKTKMSRILLGRMGYP 248
 QY 240 KEYGAVVYLSESAASYTTGASLLVDDGFTSW 272
 DB 249 DDLIGAVVFASD-ASNFTVGLIYIDGGVTSW 280

RESULT 9
 AC0157
 Probable short chain oxidoreductase YPO1287 [imported] - Yersinia pestis (strain CO92)
 C/Species: Yersinia pestis
 C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C/Accession: AC0157
 R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AC0157
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-271 <KUR>
 A:Cross-references: UNIPROT:O8ZGK4; GB:AL590842; PIDN:CAC90118.1; PID:g15979338; GSPDB:C
 C:Genetics:
 A:Gene: YPO1287
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 24.7%; Score 349; DB 2; Length 271;
 Best Local Similarity 34.3%; Pred. No. 3.5e-20;
 Matches 91; Conservative 44; Mismatches 102; Indels 28; Gaps 7;
 QY 12 GHUTVTGACGGLAELIKGLAYGSDIALDIDQEKTAQAQEHKATAE 67
 DB 29 GGRALVTAGGCGGAIAGLAATGAEVICTDISERRAAATQALNAKGVNRAEGL-- 85

QY 68 EYRKMGSYACDISDSTVHKVPAQVAKDFGKLLHLVNTAGCENPFCEYPAKNAEK 126
 DB 86 -----DYVDSALIDALNAALP-----PLDVLVCNAGVITHTPAEEMTDADWDKV 129
 QY 127 VKNVLGSLVSGAFKPLIKEGIGKASVYLIGMSGAIIVDPQNVVYNSKAGVYHLA 186
 DB 130 IANVLGVFTFCGFGFRMLIAG--RGIINISISGQIVNVPOPOCHNPAASAGVHLL 187
 QY 187 KTLACEMAKNIRVNSLNPXYIGPLTKVINGBELYNNISGIPQGRSEPEYIGAV 246
 DB 188 KSLAVEMATGVRVNAVAFYIETPLIG-LTSQPGVSRHMDMTBMGRLSPEHIAASV 246
 QY 247 LYLSESAASYTTGASLLVDDGFTS 271
 DB 247 QPLASE-ASSLITGSIITADAGYTS 270

RESULT 10
 D70635
 Hypothetical protein Rv1928c - Mycobacterium tuberculosis (strain H37RV)
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C/Accession: D70635
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Suleiron, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:9825987; PMID:9634220
 A:Accession: D70635
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-255 <COL>
 A:Cross-references: UNIPROT:P95286; GB:Z8498; GB:AL123456; NID:g3261701; PIDN:CAB06498.1
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv1928c
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 F/12-194/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 24.3%; Score 344; DB 2; Length 255;
 Best Local Similarity 33.7%; Pred. No. 8.1e-20;
 Matches 90; Conservative 45; Mismatches 108; Indels 24; Gaps 8;
 QY 8 FRPDGLTIVTGACGGLAELIKGLAYGSDIALDIDQEKTAQAQEHKATAE 67
 DB 7 FDLGKALILITGASTGIGKRVALLAYEAGVAT-----AARHLDLEKL-AD 53
 QY 68 EYRKMGS-----YACDISDSTVHKVPAQVAKDFGKLLHLVNTAGCENPFCEYPAKNA 123
 DB 54 EIGTSGKRVVPCDVSOHQOVTSMLDQVTAELGIDIAVCN-AGIITVTMLMPLPEEF 112
 QY 124 EKMKVNLGSLVYSQAFKPLIKEGIGASVVLIGMSGAIIVDPQNVVYNSKAGV 183
 DB 113 QRLONTVNTGFLTRQAARAKMVRQG-QGGVITNTASGHIIVPQOVSHYCSKAATV 171
 QY 184 HLAKTACEMAKNIRVNSLNPXYIGPLTKVINGBELYNNISGIPQGRSEPEY 243
 DB 172 HLTAMAVELAPKIRVNSVPGYI---LLE-LVEPTVEYPLMEPKPLGLRLRPDELA 227
 QY 244 GAVVYLSESAASYTTGASLLVDDGFT 270
 DB 228 GUYIYLAEE-ASSVMTGSDIYIDGGYT 253

RESULT 11
 D72377
 Oxidoreductase, short chain dehydrogenase/reductase family - Thermotoga maritima (strain
 C/Species: Thermotoga maritima
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C/Accession: D72377
 R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-255 <GRI>

A:Cross-references: UNIPROT:Q19246; EMBL:U41749; NID:g1118144; PID:g1118145; PIDN:AAB524

A:Experimental source: strain Bristol N2; clone F09E10

C:Genetics:

A:Gene: CESP:F09E10.3

A:Map position: X

A:Insertions: 5/2; 101/3; 126/3; 184/3

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 22.0%; Score 312; DB 2; Length 255;

Best Local Similarity 33.2%; Pred. No. 2.7e-17;

Matches 85; Conservative 47; Mismatches 100; Indels 24; Gaps 7;

QY 18 TGACGGALAEALIKGLAAGSDIALDIDDEKTAQAQAEYHKATBELKKEVP---KMG 74

DB 20 SGASGIGKAIQTLAKGARVAVADLDSGNA---TAKALPASQSHS 66

QY 75 YACDISDSTVHKVPAQVAKDFGKLPALVNTAGYCNPCEDYPAKNAERKVKVNLGS 134

DB 67 FACDVSNDSV-KGLSEHVKSIG--TPSLVNCAGITKSTLLKMQEQMDSVIKVNLTGV 124

QY 135 LVYSQAFKAPLIKEGIKGASVVLIGMSGALVNDPQNOVYTNMSKAGVIHLAKTLACEWA 194

DB 125 FHVSGAFVKASVDNNHPLSIINVSIVGKMGNGQTN--YAATKAGVIGFTKSAKELA 182

QY 195 KYNRVNLSNPGYIYGLTKVINGNEELYNRMISGIPQRMSEPEKVIYGAIVLLSESA 254

DB 183 KKNRVNVLPGFIPKTPTEAM--PTVLALICKGIPMRGEMETANSVLYLAD-L 238

QY 255 ASYTTGASLVDGGFT 270

DB 239 SSYVTGATLEVTGFS 254

RESULT 15

AH2362

hypothetical protein alr4456 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AH2362

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuguhi, Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2362

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-267 <KUP>

A:Cross-references: UNIPROT:Q8YNV5; GB:BA000019; PIDN:BAW6155.1; PID:g17133592; GSPDB:C

A:Experimental source: strain PCC 7120

A:Gene: alr4456

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 21.6%; Score 305.5; DB 2; Length 267;

Best Local Similarity 30.7%; Pred. No. 9.5e-17;

Matches 80; Conservative 55; Mismatches 109; Indels 17; Gaps 7;

QY 12 GHLTVTACGCGALAEALIKGLAAGSDIALDIDDEKTAQAQAEYHKATBELKKEVPK 71

DB 22 GKALITGKTKGIGLAIQAEFLALGAEVIVARNAEAIQOMKAMH-----SAGK 71

QY 72 MGSYACDISDSTVHKVPAQVAKDFGKLPALVNTAGYCNPCEDYPAKNAERKVKVNL 131

DB 72 VHGAVADVSTSGRQGMALYVSKTPELDI-LVNNVGTWIRKKATDYTEEFPAIFQINL 130

QY 132 LGSLYVSGAFKAPLIKEGIKGASVVLIGMSGALVNDPQNOVYTNMSKAGVIHLAKTLAC 191

DB 131 TSIFELSRIF-YPLLEKTS-KNSSIVIGSVAGLI--SVRTGAPYGMTKAALVOLTRSLAV 186

QY 192 EWAQYNIRVNSLNPGYIYGLTKVINGNEELYNRMISGIPQRMSEPEKVIYGAIVLLS 251

DB 187 EWDGDIRVNAIAFWFIQTLTEPLIN-NPETLSAVLSRTPMKRVGQPEE-VASLTAFLC 244

QY 252 ESAASYTTGASLVDGGFTSW 272

DB 245 MPTASYTTGQCIADVGGFIAP 265

Search completed: July 9, 2005, 10:31:43

Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Inc.

OM protein - protein search, using sw model

Run on: July 9, 2005, 10:19:36 ; Search time 178 Seconds

(without alignments)
782.503 Million cell updates/sec

Title: US-10-720-018-2

Perfect score: 1415

Sequence: 1 MTDIYPTFRFDGHLITVGA.....SASVYTGASLLVDGFTSW 272

Scoring table: BLOSUM62

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trcbl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 1415 | 100.0 | 272 | Q70FD1 | Q70FD1 ambrosiozym |
| 2 | 800 | 56.5 | 281 | ARDH_CANAL | P43066 candida alb |
| 3 | 798.5 | 56.4 | 278 | ARDH_PICST | P50167 pichia stip |
| 4 | 797 | 56.3 | 282 | 1 ARDH_CANTR | P50166 candida tro |
| 5 | 769.5 | 54.4 | 278 | 2 O6BY61 | O6BY61 debaryomyce |
| 6 | 659.5 | 47.3 | 314 | 2 O6CX46 | O6CX46 kluyveromyce |
| 7 | 528.5 | 37.3 | 343 | 2 O6C367 | O6C367 yarrowia li |
| 8 | 510.5 | 36.1 | 368 | 2 O6T5L8 | O6T5L8 emericella |
| 9 | 501 | 35.4 | 371 | 2 O7SE47 | O7SE47 neurospora |
| 10 | 465 | 32.9 | 291 | 2 O6C6W8 | O6C6W8 yarrowia li |
| 11 | 407 | 28.8 | 259 | 2 O6ESR1 | O6ESR1 manheimia |
| 12 | 402.5 | 28.4 | 257 | 2 O9WYD3 | O9WYD3 thermotoga |
| 13 | 399.5 | 28.2 | 291 | 2 O6BN16 | O6BN16 debaryomyce |
| 14 | 393 | 27.8 | 266 | 1 DCKR_TRIRE | O8H150 trichoderma |
| 15 | 392.5 | 27.7 | 257 | 2 O89Q45 | O89Q45 bradyrhizob |
| 16 | 389.5 | 27.5 | 364 | 2 O7RZD4 | O7RZD4 neurospora |
| 17 | 383.5 | 27.1 | 279 | 2 O6BQ25 | O6BQ25 debaryomyce |
| 18 | 382 | 27.0 | 281 | 1 SOUT_CANAL | P87219 candida alb |
| 19 | 381 | 26.9 | 256 | 2 O92MP6 | O92MP6 rhizobium m |
| 20 | 378 | 26.7 | 283 | 2 O9C4B3 | O9C4B3 candida mag |
| 21 | 377.5 | 26.7 | 257 | 2 O92MR6 | O92MR6 rhizobium m |
| 22 | 376 | 26.6 | 262 | 2 O8UB64 | O8UB64 agrobacteri |
| 23 | 374 | 26.4 | 280 | 2 O6CM06 | O6CM06 kluyveromyce |
| 24 | 368.5 | 26.0 | 280 | 1 SOUT_CANAL | P87218 candida alb |
| 25 | 361 | 25.5 | 255 | 2 O9Y6Z9 | O9Y6Z9 schizosacch |
| 26 | 360.5 | 25.5 | 265 | 2 O8E2P3 | O8E2P3 gibberella |
| 27 | 359 | 25.4 | 340 | 2 O7SDB3 | O7SDB3 neurospora |
| 28 | 358.5 | 25.3 | 266 | 2 O96W29 | O96W29 cladosporiu |
| 29 | 358.5 | 25.3 | 281 | 2 O29061 | O29061 archaeoglob |
| 30 | 357 | 25.2 | 254 | 2 O8B3G6 | O8B3G6 pseudomonas |
| 31 | 356 | 25.2 | 271 | 2 O66CT3 | O66CT3 yersinia ps |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------|
| 32 | 354.5 | 25.1 | 266 | 2 | O8J231 | alternaria |
| 33 | 350 | 24.7 | 271 | 2 | O7E2T4 | neurospora |
| 34 | 349 | 24.7 | 271 | 2 | O8ZGK4 | yersinia pe |
| 35 | 348.5 | 24.6 | 255 | 2 | O8COE7 | staphylococ |
| 36 | 348.5 | 24.6 | 257 | 2 | O8FZ65 | bruceella su |
| 37 | 347.5 | 24.6 | 259 | 2 | O6BN65 | debaryomyce |
| 38 | 347 | 24.5 | 256 | 2 | O92MR3 | rhizobium m |
| 39 | 345 | 24.4 | 273 | 2 | O89H04 | bradyrhizob |
| 40 | 344 | 24.3 | 255 | 2 | P95286 | mycobacteri |
| 41 | 344 | 24.3 | 255 | 2 | O7TZ97 | mycobacteri |
| 42 | 343 | 24.2 | 268 | 2 | O986J1 | rhizobium l |
| 43 | 341.5 | 24.1 | 256 | 2 | O7JZf4 | mycobacteri |
| 44 | 340.5 | 24.1 | 278 | 2 | O6CE59 | yarrowia li |
| 45 | 339 | 24.0 | 285 | 2 | O6BQ28 | debaryomyce |

ALIGNMENTS

| RESULT 1 | | | | | | | | | |
|--|---|------------------------------------|---------------------------------------|-----|-----|--|--|--|--|
| ID | Q70FD1 | PRELIMINARY; | PRT; | 272 | AA. | | | | |
| AC | Q70FD1 | | | | | | | | |
| DT | 05-JUL-2004 (TRENBLrel. 27, Created) | | | | | | | | |
| DT | 05-JUL-2004 (TRENBLrel. 27, Last sequence update) | | | | | | | | |
| DT | 05-JUL-2004 (TRENBLrel. 27, Last annotation update) | | | | | | | | |
| DE | NADH L-xyulose reductase (Fragment). | | | | | | | | |
| GN | Name=alx1; | | | | | | | | |
| OS | Ambrosiozyma monospora. | | | | | | | | |
| OC | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; | | | | | | | | |
| OC | Saccharomycetales; Saccharomycopsidaceae; Ambrosiozyma. | | | | | | | | |
| OX | NCBI_TaxID=43982; | | | | | | | | |
| RN | [1] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RC | STRAIN=NRRL Y-1484; | | | | | | | | |
| RX | PubMed=14736891; DOI=10.1074/jbc.M12533200; | | | | | | | | |
| RA | Verho R., Putkonen M., Londeborough J., Penttilae M., Richard P.; | | | | | | | | |
| RT | "A novel NADH linked L-xyulose reductase in the L-arabinose catabolic | | | | | | | | |
| RT | pathway of yeast." | | | | | | | | |
| RL | J. Biol. Chem. 279:14746-14751 (2004). | | | | | | | | |
| DR | EMBL; AJ583159; CAB47547.1; -. | | | | | | | | |
| DR | HSSP; P50162; IAB1. | | | | | | | | |
| DR | GO; GO:0016491; F:oxidoreductase activity; IEA. | | | | | | | | |
| DR | GO; GO:0008152; P:metabolism; IEA. | | | | | | | | |
| DR | InterPro; IPR002198; ADH_short. | | | | | | | | |
| DR | InterPro; IPR002347; Adh_short_C2. | | | | | | | | |
| DR | InterPro; IPR002110; ANK. | | | | | | | | |
| DR | Pfam; PF00106; adh_short; 1. | | | | | | | | |
| DR | PRINTS; PRO1415; ANKYRIN. | | | | | | | | |
| DR | PRINTS; PRO0081; GDERDH. | | | | | | | | |
| DR | PROSITE; PS00061; ADH_SHORT; UNKNOWN_1. | | | | | | | | |
| FT | NON TER 272 | | | | | | | | |
| SO | SEQUENCE 272 AA; 29636 MW; 3338DB95ECC473D CRC64; | | | | | | | | |
| Query Match 100.0%; Score 1415; DB 2; Length 272; | | | | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 7.3e-107; | | | | | | | | | |
| Matches 272; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| Qy | 1 | MTDIYPTFRFDGHLITVGA | CGGIAELIKGLAYGSDIALLDIDDKTKAKQAEYHKYA | 60 | | | | | |
| Db | 1 | MTDIYPTFRFDGHLITVGA | CGGIAELIKGLAYGSDIALLDIDDKTKAKQAEYHKYA | 60 | | | | | |
| Qy | 61 | TEELKKEVPKMSYACDISDPTVHKVFAQVAKD | RGKLPPLHVNTPAGYENPFCEYPA | 120 | | | | | |
| Db | 61 | TEELKKEVPKMSYACDISDPTVHKVFAQVAKD | RGKLPPLHVNTPAGYENPFCEYPA | 120 | | | | | |
| Qy | 121 | KNAEKVKVNLGSLVYSAFAKPLIKGIGASVVL | IGSMGAIYNDPQNVVYNNSKA | 180 | | | | | |
| Db | 121 | KNAEKVKVNLGSLVYSAFAKPLIKGIGASVVL | IGSMGAIYNDPQNVVYNNSKA | 180 | | | | | |
| Qy | 181 | GVVHAKTLACFAKYNIRNSLNPQYITGLTKVYNG | BEELYNRISGIPQORNSBP | 240 | | | | | |
| Db | 181 | GVVHAKTLACFAKYNIRNSLNPQYITGLTKVYNG | BEELYNRISGIPQORNSBP | 240 | | | | | |

OY 241 EYIGAVLYLSESAASYTTGASLLVDGGFTSM 272
 DB 241 EYIGAVLYLSESAASYTTGASLLVDGGFTSM 272

RESULT 2

ARDH CANAL ID ARDH CANAL STANDARD; PRT; 281 AA.
 AC P43066; 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE D-arabinol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250) (ARDH).
 GN Name-ARDH; Synonyms-ARDH;
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=54476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MO-1;
 RX MEDLINE=94012494; PubMed=8407803;
 RA Wong B., Murray J.S., Castellanos M., Croen K.D.;
 RT "D-arabinol metabolism in Candida albicans: studies of the
 biosynthetic pathway and the gene that encodes NAD-dependent D-
 arabinol dehydrogenase.";
 RL J. Bacteriol. 175:6314-6320(1993).
 CC -1- CATALYTIC ACTIVITY: D-arabinol + NAD(+) = D-ribulose + NADH.
 CC -1- PATHWAY: D-arabinol biosynthesis; last step.
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L16227; AAC37430.1; -.
 DR HSP; Q9ZFY9; 1FK8.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short_1.
 DR PRINTS; PR00081; GDRDH.
 DR PRINTS; PR00080; SDRPAMILY.
 DR TIGRFAMs; TIGR01830; 3oxo_ACP_reduc; 1.
 DR TIGRFAMs; TIGR01831; fadG_rel; 1.
 DR TIGRFAMs; TIGR01832; kdud; 1.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW NAD; Oxidoreductase.
 FT NP_BIND 25
 FT ACT_SITE 184 184 NAD (By similarity).
 FT BINDING 169 169 Proton acceptor (By similarity).
 FT BINDING 169 169 Substrate (By similarity).
 SQ SEQUENCE 281 AA; 30643 MW; B302A6411591F892 CRC64;

Query Match 56.5%; Score 800; DB 1; Length 281;
 Best Local Similarity 55.7%; Pred. No. 8e-57;
 Matches 151; Conservative 55; Mismatches 61; Indels 4; Gaps 2;

OY 5 IPTFRFDGLTIVTGACGGLAEALIKGLAYGSDIALIDDEKTAQAQAEYKATTEEL 64
 DB 12 VSFRLDGLVITIGSGGLAAVSRALAKGADVALLVDMILERTQAARDVLQMGEEQM 71
 OY 65 KKK---EVRKMGSYACDISDSDTVHKVPAQVAKDFEKLPLHLVNTAGYCEPFCEDEYPAK 121
 DB 72 KGVSEPIQGVSAWMSCNIGDAEAVDLTFALINEHKGISVYVNTAGYENFPAAEYPAK 131
 OY 122 NAEKMYKVNLLGSLYVSOAPAKFLKEGKASVLLIGMSGATVNDPQNVVYVYNSKAG 181
 DB 132 NENLMKVNGLGFVYVSOAFARPLIONMTG--SIILIGSGITVNDPQNVYVYNSKAG 190

OY 182 VVHAKTLACEMAKYNIIRVNSLNPXYIGPLTKVINGNEELVNRWISGIPQRMSEPYE 241
 DB 191 VVHAKSLACEMAKYNIIRVNTLSPGYILPLFTRNVISGHTKMEWESKIPMRMAEPYE 250
 OY 242 YIGAVLYLSESAASYTTGASLLVDGGFTSM 272
 DB 251 FVGSILYLAESASASYTTGHNLVVDGGEBCW 281

RESULT 3

ARDH PICST ID ARDH PICST STANDARD; PRT; 278 AA.
 AC P50167;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE D-arabinol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250) (ARDH).
 GN Name-ARDH;
 OS Pichia stipitis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4924;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 58785 / CBS 6054;
 RX MEDLINE=96090133; PubMed=7483848;
 RA Hallborn J., Walfridsson M., Penttilae M., Keraenen S.,
 RA Hann-Haegderl B.;
 RT "A short-chain dehydrogenase gene from Pichia stipitis having D-
 arabinol dehydrogenase activity.";
 RL Yeast 11:839-847(1995).
 CC -1- CATALYTIC ACTIVITY: D-arabinol + NAD(+) = D-ribulose + NADH.
 CC -1- PATHWAY: D-arabinol biosynthesis; last step.
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z46866; CAA86939.1; -.
 DR PIR; S57351; S57351.
 DR HSP; Q9ZFY9; 1FK8.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short_1.
 DR PRINTS; PR00081; GDRDH.
 DR PRINTS; PR00080; SDRPAMILY.
 DR TIGRFAMs; TIGR01830; 3oxo_ACP_reduc; 1.
 DR TIGRFAMs; TIGR01831; fadG_rel; 1.
 DR TIGRFAMs; TIGR01832; kdud; 1.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW NAD; Oxidoreductase.
 FT NP_BIND 22
 FT ACT_SITE 181 181 NAD (By similarity).
 FT BINDING 166 166 Proton acceptor (By similarity).
 FT BINDING 166 166 Substrate (By similarity).
 SQ SEQUENCE 278 AA; 30003 MW; 36869165F23964Fe CRC64;

Query Match 56.4%; Score 798.5; DB 1; Length 278;
 Best Local Similarity 54.3%; Pred. No. 1e-56;
 Matches 151; Conservative 54; Mismatches 64; Indels 9; Gaps 3;

OY 3 DY-----IPTFRFDGLTIVTGACGGLAEALIKGLAYGSDIALIDDEKTAQAQAEYH 57
 DB 2 DYANVAVVNFRLDGLALITIGSGGLAAVSRALAKGADVALLVDMILERTKSAKAEVL 61
 OY 58 KATATELKLKEVPPKKG---STACDISDSDTVHKVPAQVAKDFEKLPLHLVNTAGYCEPF 114

DB 62 GNGEETLKGEHSAIQGVAMSCNIGDAVDATESSINEHCKIADLLINTAGYCENP 121

QY 115 CEDYPKAKNAERKVKNNLGSLLVYSOAFAPKLKEGIGASVVLIGSMGAIYNDPONQV 174
 122 AATYPTATNAMESIMKVGSGSPVYSGFAPPLIQNNLRG-SILLIGSMGTYNDPQCM 180

DB 175 YMSKAGVHHLAKTLACCAKYNIRVNSLNPGLYIPLTKVYINGNEELYNMISGIPQO 234
 181 YMSKAGVHHLVRSIACCAKYNIRVNTLSPGYILPLTRNVI SGHTEKEMKAMESKIPMK 240

QY 235 RMSPEPEYIGAVLYLSESAASYTTGASLLVNDGGFTSW 272
 241 RMAEPKEFVGSLIYLASSTASSTTGHNLVVDGGYECW 278

DB

RESULT 4
 ARDH_CANTR ID STANDARD; PRT; 282 AA.
 AC P50166;
 DT 01-OCT-1996 (rel. 34, Created)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 25-OCT-2004 (rel. 45, Last annotation update)
 DE D-arabinitol 2-dehydrogenase [ribulose forming] (EC 1.1.1.250) (ARDH).
 GN Name:ARD;
 OS Candida tropicalis (Yeast).
 OC Buxaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 ON NCBI_TaxId=54482;
 RX STRAIN=ATCC 750;
 MEDLINE=95212917; PubMed=7698655; DOI=10.1016/0378-1119(94)00900-D;
 RA Murray J.S., Wong M.L., Miyada C.G., Switchenko A.C., Goodman T.C.,
 Wong B.;
 RT "Isolation, characterization and expression of the gene that encodes
 D-arabinitol dehydrogenase in Candida tropicalis.";
 RL Gene 155:123-128(1995).
 RN (12)
 RP CHARACTERIZATION.
 RX MEDLINE=94071892; PubMed=8250887;
 RA Quong M.W., Miyada C.G., Switchenko A.C., Goodman T.C.;
 RT "Identification, purification, and characterization of a D-arabinitol-
 specific dehydrogenase from Candida tropicalis.";
 RL Biochem. Biophys. Res. Commun. 196:1323-1329(1993).
 CC -1- CATALYTIC ACTIVITY: D-arabinitol + NAD(+) = D-ribulose + NADH.
 CC -1- PATHWAY: D-arabinitol biosynthesis; last step.
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U00675; AAA6355.1; -
 DR PIR: J04041; J04041.
 DR HSSP: Q9ZFY9; 1FK8.
 DR InterPro: IPR002198; Adh_short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam: PF00106; adh_short_1.
 DR PRINTS: PR00081; GDRDH.
 DR PRINTS: PR00080; SDRFAMILY.
 DR TIGRFAMs: TIGR01830; 3oxo_ACP_reduc; 1.
 DR TIGRFAMs: TIGR01831; fadG_tel; 1.
 DR TIGRFAMs: TIGR01832; kdud; 1.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KW Direct protein sequencing; NAD; Oxidoreductase.
 NP BIND 26 48
 FT ACT_SITE 185 185 NAD (By similarity).
 FT BINDING 170 170 Proton acceptor (By similarity).
 FT Substrate (By similarity).

SQ SEQUENCE 282 AA; 30748 MW; A82A3DA4E771EF0D CRC64;
 Query Match 56.3%; Score 797; DB 1; Length 282;
 Best Local Similarity 55.0%; Pred. No. 1.4e-56;
 Matches 149; Conservative 57; Mismatches 61; Indels 4; Gaps 2;

QY 5 IPTFRPDGLTIVTGAACGLAEALIKGLIYSGDIALDIDQKTPAAQAEYHKYATEEL 64
 13 VPSFLDGLKVIITGGSGGLSAVVRALLAKADIALIDMNLERTQQAARDVLQMGCEQM 72

DB 73 KKHESPIGVSAWAMSCNIGDAEAVELTPVALINEHKGASVYLINTAGAEENPAAEYPPK 132

QY 65 KKE---VPMKSYACDISDSDPTNKKVPAQVAKDGGKPLHLVNTNAGYCENPCEDPK 121

DB 122 NAEKVKVNLGSLVYSOAFAPKLKEGIGASVVLIGSMGAIYNDPONQVYVNSKAG 181
 133 NAEKVKVNLGSLVYSOAFAPKLQNNMTG-SILLIGSMGTYNDPQCMYVNSKAG 191

QY 182 VIHLAKTLACCAKYNIRVNSLNPGLYIPLTKVYINGNEELYNMISGIPQORMSEPK 241
 192 VIHLARSLACCAKYNIRVNTLSPGYILPLTRNVI SGHTEKEMKAMESKIPMKRAEPK 251

DB 242 YIGAVLYLSESAASYTTGASLLVNDGGFTSW 272
 252 FVGSILYILASDSASSYTTGHNLVVDGGYECW 282

QY

RESULT 5
 OGBY61 ID PRELIMINARY; PRT; 278 AA.
 AC OGBY61;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Similar to CA3288|CAARD8 Candida albicans CAARD8 D-arabinitol
 dehydrogenase.
 GN ORFNames=DEHA0A12595g;
 OS Debaryomyces hansenii CBS767.
 OC Buxaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
 ON NCBI_TaxId=284592;
 RX (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS767;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 Latouche I., de Montigny J., Marck C., Nevegilise C., Talla E.,
 Goffard N., Frangeul L., Aigle M., Anthonard V., Babour A., Barbe V.,
 Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 Boiserame A., Boyer J., Catolico L., Confanioleri F., de Darvar A.,
 Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,
 Hantrege F., Hemequin C., Jauniaux N., Joyet P., Kachouri R.,
 Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 Nicoud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
 Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 Swenne S., Tekla F., Weselowski-Louvel M., Westhof E., Wirth B.,
 Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts.";
 RL Nature 430:35-44(2004).
 RN (12)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS767;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.
 DR EMBL: CR382133; CAG84833.1; -
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0008152; F:metabolism; IEA.
 DR InterPro: IPR002198; Adh_short.
 DR InterPro: IPR002347; Adh_short_C2.

(SDR) family.
 CC EMBL; 382132; CNG77697.1; -
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; Adh_short_C2.
 DR InterPro; IPR002347; Adh_short_C2.
 DR InterPro; IPR00205; NAD_BS.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00081; GDRHDH.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 DR Oxidoreductase.
 KW OXIDOREDUCTASE.
 SQ SEQUENCE 343 AA; 37082 MW; 1E593CF7047E5585 CRC64;
 Query Match 37.3%; Score 528.5; DB 2; Length 343;
 Best Local Similarity 36.9%; Pred. No. 1.3e-34;
 Matches 125; Conservative 55; Mismatches 70; Indels 89; Gaps 10;
 QY 16 IVTGACGGLAELIKGLAYGSDIALLD--IDQKTAQKQ-ARHKYKATE----- 62
 DB 12 VITGGSGGLGIAAQLAAQASVALVDNNLPRIQPAEQLLKEMKYNKAHNNVPTPEI 71
 QY 63 -----BLKL----- 66
 DB 72 YASPTGTRKVSSTETESTTGLENESPHDITPDISLSDASDSSQSVAHDAARAHEAG 131
 QY 67 -----KEVP--KMSYACDISD---SPTVHKVPAQVAKDPGKLP-LVNTAGYCENF 113
 DB 132 IPRGKQKQPPQORISAMACDVSDVHQVSDTVKAI-----REHKKSPDLIDVNCAPFCENM 186
 QY 114 PCEDYPAKKAEMKYNKVLGSLVYSQAPKPLIKGIGKASVVLIGSMGAILVNDPQOV 173
 DB 187 TAFDPNPQVKRLDLVNLMGSSYNFATEVAKSLVLDSPG-SLILVASMGSGSLVNDPQOP 245
 QY 174 VYNNKAGVYHIAKTLACEMAKYNTRVNSLNPGYTGPITKVNINSELYNRTISGIPQ 233
 DB 246 PYNNSKAGVYHIAKSLAEMAKYNTRVNSLNPGYTGPITKVNINSELYNRTISGIPQ 305
 QY 234 QRMSEPKKEYIGAVLYLSESASVYTTGASLLVDGGFTSW 272
 DB 306 RMAEPSEFGPIVPMASD-ASSVMTGHDLDVGGGTYTW 343
 RESULT 8
 Q6T5L8 PRELIMINARY; PRT; 368 AA.
 AC Q6T5L8;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE D-arabinitol dehydrogenase (EC 1.1.1.250).
 GN Name=arbd;
 OS Emericella nidulans (Aspergillus nidulans).
 OC Bacteria; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutectales; Trichocommataceae; Emericella.
 OC NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rodriguez J.M., Ruiz-Sala P., Ugarte M., Penajaya M.A.;
 RL Submitted (OCT-2003) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=14612443; DOI=10.1074/jbc.M310055200;
 RA Rodriguez J.M., Ruiz-Sala P., Ugarte M., Penajaya M.A.;
 RT "Pungal Metabolic Model for 3-Methylcrotonyl-CoA Carboxylase
 Deficiency.";
 RL J. Biol. Chem. 279:4578-4587(2004).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.
 CC EMBL; AY434691; AAR65467.1; -
 CC EMBL; AY387591; AAR23109.1; -
 DR GO; GO:0047038; F:D-arabinitol 2-dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; Adh_short_C2.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00081; GDRHDH.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 DR Oxidoreductase.
 KW OXIDOREDUCTASE.
 SQ SEQUENCE 368 AA; 39686 MW; 67DB6F1CBE44CF62 CRC64;
 Query Match 36.1%; Score 510.5; DB 2; Length 368;
 Best Local Similarity 40.3%; Pred. No. 4e-33;
 Matches 106; Conservative 60; Mismatches 94; Indels 3; Gaps 3;
 QY 8 FRPDGLTIVGACCGGLAELIKGLAYGSDIALLDIDQKTAQKQAEYHKYATEELK 67
 DB 107 FSLERVYTLVVGARGGLVWGQGI VASGSDLAIDLDLHAEBSARKVLEQKQENPGLE 166
 QY 68 EYPKMGSYACDISDSSTVHKVPAQVAKDPGKLP-LVNTAGYCENFPCEDYPAKKAEMK 127
 DB 167 QLPNTAHYADVNSPDSVNNALSEVIAHGKID-HLVTSAGFTENFDIAISYPDRMQKLW 225
 QY 128 KVNLLGSLVYSQAPKPLIKGIGKASVVLIGSMGAILVNDPQOVVYNNKAGVYHIAK 187
 DB 226 GNVVDGYTLFATGVAKHMERKAPG-STVMIGSMGAILVNDPQAPYNAALAAVRHLAS 284
 QY 188 TLACEMAKYNTRVNSLNPGYTGPITKVNINSELYNRTISGIPQRMSEPKKEYIGAVL 247
 DB 285 SPAYEMAHGIRVNCISGVMLTALTRILDENBELRKWLSLPTGKQVPELMGAVT 344
 QY 248 YLSESASVYTTGASLLVDGGFT 270
 DB 345 FLTSD-ASKYITGADLRVDGYT 366
 RESULT 9
 Q7SE47 PRELIMINARY; PRT; 371 AA.
 AC Q7SE47;
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=NCU02128.1;
 OS Neurospora crassa.
 OC Bacteria; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OC NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., Fitzhugh W., Ma L.-O., Smirnov S., Purcell S., Reisman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Iankov P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zeller A., Schulte U.,
 RA Kohe G.O., Jedd G., Mewes W., Straben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnetre S.,
 RA Kanai M., Kamysseilis M., Mauceli E., Biele C., Rudd S., Frisman D.,
 RA Kryzstofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Varden O., Piamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbols D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
 RL Nature 0:0-0(2003).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 (SDR) family.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; AABX0100038; EAA35050.1; -

DR HSPB; P42556; 1P33.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; Adh_short_C2.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00081; GDRHD.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 DR Hypothetical protein; Oxidoreductase.
 KW SEQUENCE 371 AA; 39631 MW; 4BC6D601508DFE45 CRC64;
 SO SEQUENCE
 Query Match 35.4%; Score 501; DB 2; Length 371;
 Best Local Similarity 40.3%; Pred. No. 2, 4e-32;
 Matches 108; Conservative 57; Mismatches 97; Indels 6; Gaps 5;
 QY 5 IPTFRPFDHLITVTGACGGLAELIKGLANGSDIALIDIDQKTAQAQAE-YHKYATE 62
 DB 98 LATFSLQGVGVVTDGARGGLGVMGQGIYVSGDLAIYDLNKEE-ATKQAQNIYETPFKD 156
 QY 63 ELKLEKVPKMGSYADIDSDPTVHKVPAQVANDPKRLPLHLVNTGYCENPFCEYPAKN 122
 DB 157 YPSAKKIPKPTVTHYADVSDPASPVDACIAKEIVAEHKRID-NLVTSGAFENFPAVYVDR 215
 QY 123 AEKMKVNLGLSLVYQAQPAKPLIKEGIKGASVTLIGSMGAIYVDPQNVVYNNKSKAV 182
 DB 216 LRLKLVGVNDGYLPAATALARHLMDRNSG-SLVITIGSMGSIYVNPQQTLYNASKAV 274
 QY 183 IHLAKTLACEMAKYNIIRVNSLNGPYIYGPLTNNVINGEELYNRMISGIPQGRMSEPKY 242
 DB 275 RLHAASPANEMAKAGIRVNCISPGVWLTLVTKKIDENPDLKEKWTSLIPQGMGNPDL 334
 QY 243 ICAVLYLSESAASYTTGASLVDGFT 270
 DB 335 MGPVTFLLSD-ASRYVTGADLRVDGGYT 361
 RESULT 10
 Q6C6W8 PRELIMINARY; PRT; 291 AA.
 AC Q6C6W8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Similar to tr|Q9WVD3 Thermotoga maritima TM0297 Oxidoreductase.
 GN ORFNames=YALI0E05643g;
 OS Yarrowia lipolytica CLIB99.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxId=284591;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marc C., Neugeglise C., Talla E.,
 RA Goffard N., Franguel L., Algite M., Authourd V., Babout A., Barbe V.,
 RA Barray S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Cattolico L., Confantolieri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,
 RA Hantaye F., Hemequin C., Janniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Kozul R., Lemaitre M., Lestur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Ozras S., Ozier-Kalogeropoulos O.,
 RA Pellens S., Poctier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swemene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Bolocin-Fukuhara M., Thierry A.,
 RA Bouchier C., Gaudion J.L.,
 RA Wincker P., Souciet J.L.,
 RT "Genome evolution in yeasts."
 RL Nature 430:35-44(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIB99;

```

RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -i- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
   (SDR) family.
DR EMBL: CR382131; CAG79175.1; -.
DR GO: GO:0016491; P:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002347; Adh_short_C2.
DR Pfam: PF01065; adh_short; 1.
DR PRINTS: PR00081; GDERDH.
DR PRINTS: PR00080; SDRFAMILY.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 291 AA; 31002 MM; 9479873014C401D CRC64;

Query Match 32.9%; Score 465; DB 2; Length 291;
Best Local Similarity 36.9%; Pred. No. 1.5e-29;
Matches 103; Conservative 55; Mismatches 85; Indels 36; Gaps 7;

QY 3 DYIPFRPDGHLITVGACGGIAELIKGLAYG-SDIALDIDOE-----KTA 51
DB 36 NPLSFRIDGVAALVTGGARLGGFMAELGCVSLKRIALILDVQDGLDAIEKHKAY 95
QY 52 KOAEHYKATEELKLEVPKMSYACDISDVTYHKVFAOVADFGKLPILYNTYGYCE 111
DB 96 VQAGYK-----ADVDESVNEILDRVHDLGSDV-VNNSAGVAD 136
QY 112 NPPCEDYPAKAEMKVKVNLGLSLYVQAPKPLIKGIGIGASVYLIGMSGAILVNDPON 171
DB 137 LVHAAYEPADPRRRIYDINLNGSEFLVYQAARRHVIKGG-TGGTVFASNSGSLVWPP 195
QY 172 QVVNMYSGAYIHLAKTACEMAKXNTRVNSLNFPGYIGPLTKVINGNEELYNRWISGI 231
DB 196 GSAYNASRAAYHSHKSLAAMAVAHNIRKNSISFGYVDIALNR-----AYNTLFEWMDRT 251
QY 232 POORMSEPEKEYIGAVLYLISEAASYYTGASLVDGFT 270
DB 252 PLGRLDPELDLGACTIYASD-ASSYVYTGSDIITIDGYT 289

RESULT 11
Q6SSP1 PRELIMINARY; PRT; 259 AA.
AC Q6SSP1;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DB Fabg protein.
GN Name=fabg; ORFNames=MS1412;
OS Mannheimia succiniciproducens MBEL55E.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxId=221988;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=MBEL55E.
RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
RA Kim C.H., Jeong H., Hur C.-G., Kim U.G.;
RA "The genome sequence of the capnophilic rumen bacterium Mannheimia
RA succiniciproducens."
RL Nat. Biotechnol. 0:0-0(2004).
DR EMBL: AB016827; AAU38019.1; -.
SQ SEQUENCE 259 AA; 27552 MM; B46BF478B3F5923C CRC64;

Query Match 28.8%; Score 407; DB 2; Length 259;
Best Local Similarity 37.4%; Pred. No. 6.7e-25;
Matches 99; Conservative 44; Mismatches 104; Indels 18; Gaps 6;

QY 9 RPDGHLITVGACGGIAELIKGLAYGSDIALDID--QETPAKQAEYHXYATEELK 65
DB 8 KLTKGTAFTVGGARGIGKSVLAIFAQAGANVVIAFDIAEAKTIA-----EIA 56

```

QY 66 LKEVPKMSYACDISDSTVHKVFAQVAKDREGKLPPLHVTAGYCENFCEDYPAKMAEK 125
 DB 57 KEEGVKSAJAVOTDVTDOASVNHLMVDVIKQFCKLIDAFCN-AGICINPAEEMSYEOMLK 115
 QY 126 MKVKNVILGSLYVQAFAKPLIKEGIKGASVVLIGSMGAIIVNDPONQVYNNMSKAGVITL 185
 DB 116 VINVLNGLVFLQAQAGKLMIEGQ-TGGSIIINTASMAHIVVPPQCAVYNSKAGVITL 174
 QY 186 AKTLACBAKXNIRVNSLNPGYITPPLTKVNLNGEELYNRMISGIPQORMSEPREYITGA 245
 DB 175 TKSLAIEWAKNIRVNSISPGYIGTELTLN-SKDQPLIKENWANAFLHRLCKPRELDSI 233
 QY 246 VLYLSESASATYTGASILLVDGFT 270
 DB 234 CVYLAGDT-SSPTGADPVDGAF 257
 RESULT 12
 Q9WYD3 PRELIMINARY; PRT; 257 AA.
 ID Q9WYD3
 AC Q9WYD3;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 12, Last annotation update)
 DE Oxidoreductase, short chain dehydrogenase/reductase family.
 GN OrderedLocustNames=TW0297;
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 NC NCB1_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
 RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey B.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L.A., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,
 RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RL Nature 399:323-329(1999).
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 CC (SDR) family.
 CC EMBL: AB001711; AAD3385.1; -
 DR PIR: A72395; A72395.
 DR HSSP: O9ZFY9; 1FTH.
 DR TIGR: TM0297;
 DR GO: GO:0016491; P:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PRO0081; GDRDH.
 DR PROSITE: PRO0080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 DR Complete proteome; Oxidoreductase.
 KM SEQUENCE 257 AA; 27886 MW; 724484FF3EDF5EAC CRC64;
 QY Query Match 28.4%; Score 402.5; DB 2; Length 257;
 QY Best Local Similarity 36.9%; Pred. No. 1.5e-24;
 DB Matches 101; Conservative 50; Mismatches 90; Indels 33; Gaps 11;
 QY 8 FRFGHLLTVYAGCGLAALIKGLAYGSDIALLDIOEKTAQAEYHXYATEELK 67
 DB 8 FSLKRAVAVTGGGOGIGKALIAQALAAAGAVLLINDINER-TARRTYE----- 54
 QY 68 EVPPKGS---YACDISDSTVHKVFAQVAKDREGKLPPLHVTAG---YCENFCEDYPA 120
 DB 55 EIKKSGEADRYVDVTKEDPCFAGVKKALDRWGLDIG-VNNAIGIDWCE---AENVPV 110
 QY 121 KMAEKMKVNLGSLYVQAFAKPLIKEGIKGASVVLIGSMGAIIVNDPONQVYNNMSKA 180

DB 111 EKKKVVIVDNLVG-VFLSAKAEFLAMKER-KYKGTINIASGHIYVKPQKOTAYNASKA 168
 QY 181 GVHILAKTIACEMAKYXNIRVNSLNPGYITGP--TKVYINGNEELYNRMISGIPQORMSE 238
 DB 169 GVHILTRSLAEMAPYGRVNSISRGYIRTLISPNV-----KDLVPLMDMPLIGRIGE 224
 QY 239 PREYIGAVLYLSESASATYTGASILLVDGFTSW 272
 DB 225 VDDLGAIAIFLAS-PASDYMTHGHDIVIDGYTWM 257
 RESULT 13
 Q6BN16 PRELIMINARY; PRT; 291 AA.
 ID Q6BN16
 AC Q6BN16;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Similar to YAL10B16192g Yarrowia lipolytica.
 GN ORFNames=DEHA0F01177g;
 OS Debaryomyces hanseni CBS767.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Debaryomycetes.
 NC NCB1_TaxID=284392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS767;
 RG GenomLevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
 RA Gottard N., Frangeul L., Aigle M., Anthouard V., Babour A.,
 RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boursane A., Boyer J., Catellacci L., Confiantieri F., de Daruvar A.,
 RA Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantrepe F., Hennequin C., Janniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,
 RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swenne S.D., Tekala F., Wesolowski-Louvel M., Westhof B., Wirth B.,
 RA Zenlin-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weisenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts."
 RL Nature 430:35-44(2004).
 CC [2]
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
 CC (SDR) family.
 CC EMBL: CR382138; CAG88708.1; -
 DR GO: GO:0016491; P:oxidoreductase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR002198; ADH_short.
 DR InterPro: IPR002347; Adh_short_C2.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PRO0081; GDRDH.
 DR PROSITE: PRO0080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 DR Oxidoreductase.
 KM SEQUENCE 291 AA; 31685 MW; 7386F67E7D8E9BD CRC64;
 QY Query Match 28.2%; Score 399.5; DB 2; Length 291;
 QY Best Local Similarity 36.8%; Pred. No. 3.2e-24;
 DB Matches 100; Conservative 50; Mismatches 96; Indels 27; Gaps 10;
 QY 5 IPRFPGHLLTVYAGCGLAALIKGLAYGSDIALLDIOEKTAQAEYHXYATEEL 64
 DB 35 IDLEFLKGVAVTGGGOGIGYALCEAVYAGAVLAWDSIDSTVATKLE----- 85
 QY 65 KLEKVPKMSYACDISDSTVHKVFAQVAKDREGKLPPLHVTAGYCENFC-----EDY 118

```
Db 86 KLVGV-KLRSYCVDTSDKEVEQTNNIVEDFDIDIFVAN-AGI--NIPVGLINEENS 141
Qy 119 PAKNAEMKMYKVNILGSLYSGAPFKPLIK--EGIKASVVLIGMSGALVNDPQNVVN 176
Db 142 NDKRWMLVMDINILNGVYCAKNGKVSRRKEGRKG-SMITGSMGSHIINTPVHQAAYN 200
Qy 177 MSKAGVHLKTLACEMAKNIRVNSLNPYIYGPLTKNVINGNEELYNRMISGIPQORM 236
Db 201 ASKASVHFPAKSLIEVDVF-ARVNSISPGVNSGINDHLPSNTRK---RWSSTIPMGRE 256
Qy 237 SEPEKRYIGAVLYLSESASAYTTGASLLVDGFE 269
Db 257 GLPELVGAYLYLASD-ASTYTTAADITVDGWM 288

RESULT 14
DCXR_TRIPE STANDARD; PRT; 266 AA.
AC O8WK50;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE L-xylulose reductase (EC 1.1.1.10) (XR).
GN Name=Lxrl;
OS Trichoderma reesei (Hypocrea jecorina).
OC Baktariota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RP SEQUENCE FROM N.A., AND ENZYME ACTIVITY.
RX MEDLINE=22005794; PubMed=12009906; DOI=10.1021/bi025529i;
RA Richard P., Putkonen M., Vasaenen R., Londestorouh J.,
RA Penttilae M.;
RT "The missing link in the fungal L-arabinose catabolic pathway,
RT identification of the L-xylulose reductase gene.";
RL Biochemistry 41:6432-6437(2002).
CC -I- FUNCTION: Catalyzes the NADPH-dependent reduction of L-xylulose,
CC D-xylulose, D-fructose, and L-sorbose, with the highest affinity
CC for L-xylulose.
CC -I- CATALYTIC ACTIVITY: Xylitol + NADP(+) = L-xylulose + NADPH.
CC -I- PATHWAY: L-arabinose pathway; third step.
CC -I- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF375616; AAM20896.1; -.
DR HSSP; Q9ZFY9; 1FK8.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRHDL.
DR PRINTS; PR00080; SDRFAMILY.
DR TIGRFAMS; TIGR01830; 3oxo_ACP_reduc; 1.
DR TIGRFAMS; TIGR01831; fabG_rel; 1.
DR TIGRFAMS; TIGR01832; kdud_1.
DR PROSITE; PS00061; ADH_SHORT; PALSE_NEG.
KW Carbohydrate metabolism; NADP; Oxidoreductase; Xylulose metabolism.
FT ACT_SITE 174 174 NADP (By similarity).
FT ACT_SITE 178 178 Proton acceptor (By similarity).
FT BINDING 159 159 Substrate (By similarity).
SQ SEQUENCE 266 AA; 28478 MW; 1CF56334DA86F109 CRC64;

Query Match 27.8%; Score 393; DB 1; Length 266;
Best Local Similarity 36.4%; Pred. No. 9; 6e-24;
Matches 102; Conservative 51; Mismatches 103; Indels 24; Gaps 8;
```

```
Qy 1 MTDIPT-----FPFDGHLITVTGACG--GLAELIKGLAYSGDIALLDIDDEKTA 51
Db 1 MPQPVPTANRLDLFLSLKGVVVVTSAGBGRKGIGI BAARGCAMGADLAITYSSRKEGAE 60
Qy 52 KQAEHYKATVEELKLKVEPVMGSIYACDISPDVTHKVFPAQVADPFCKLPLHLVNTGYCE 111
Db 61 KNA-----EBLTKREVGVKVKYKNSDYNVERFVNQVSDPFKIDAFIAN-AGATA 112
Qy 112 NFPCEDVPKNAEMKMYKVNILGSLYSGAPFKPLIKEGIKASVVLIGMSGALVNDPQN 171
Db 113 NSGVDSASDMDVHVIOVDSGTAYCAKAVGARFKQG--HGSVLVITASMSGVANYPOE 170
Qy 172 QVYVNSKAGVYHLAKTLACEMAKNIRVNSLNPYIYGPLTKNVINGNEELYNRMISG 231
Db 171 QTSYVNAKAGCICHLARSLANEMWDF-ARVNSISPGYIDTGLSDFIDEKTOEL---WRSMI 226
Qy 232 PQORMSEPEKRYIGAVLYLSESASAYTTGASLLVDGFTS 271
Db 227 PMGRNGDAKELKGAAYLVSD-ASSYTTGADIVIDGGYTT 265
```

RESULT 15

Q89QA5 PRELIMINARY; PRT; 257 AA.

```
ID Q89QA5;
AC Q89QA5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Oxidoreductase.
GN OrderedLocustNames=bl3225;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RP SEQUENCE FROM N.A.
RX STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneo T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Saramoto S., Watanabe A., Idegawa K., Iriuchi M., Kawashima K.,
RA Kohara M., Matsuno M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -I- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC EMBL; AP005946; BAC48490.1; -.
DR HSSP; Q9ZFY9; 1FJH.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDRHDL.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 257 AA; 27403 MW; A49BE3BF7F21363D CRC64;
```

```
Query Match 27.7%; Score 392.5; DB 2; Length 257;
Best Local Similarity 36.1%; Pred. No. 1e-23;
Matches 99; Conservative 45; Mismatches 107; Indels 23; Gaps 8;

Qy 4 YIFPRFGHLITVTGACGGLAEALIKGLAYSGDIALLDIDDEKTAQKQAEHYKATVEE 63
Db 2 YLEKPKLDKGKTAFTGGGQGLACAEALABAGRVVIGDDSKVADSAKA-----S 53
Qy 64 LKLK-----EVPKMGSYACDISPDVTHKVFPAQVADPFCKLPLHLVNTGYCEN--FPCEYD 118
Db 54 LKAKGFDAETALM-----DVTITRRVAEVANDVARRGKVDI-LVNNAGIGARSETPATLV 107
```

Qy 119 PAKNAEKWVNLGSLVYSOAFAPKPLKEGIKGASVYLIGSMGAIYNDPONOVYNNMS 178
 Db 108 TDEHMLNVIDVNLNGTFWCCREFGKMLK-AQSGAIVNVGSMGPIVNPQEQCFYNAS 165
 Qy 179 KAGVHLAKTLACENAKYNIIRVNSLNPGYIYGPLTKENVINGNEELYNRWISGIPOORMSE 238
 Db 166 KAGVHLTKSLAEEKGARGIRVNAVAPTYIETPLNA-FVKSINPKYTDAMIGTPMARMQ 224
 Qy 239 PREYIGAVLYLISESASYYTGTASLLVDGGFTSW 272
 Db 225 VEIRASVVLFLSSB-AASIMTGSIVLVDGGYTCW 257

Search completed: July 9, 2005, 10:30:58
 Job time : 181 secs

This Page Blank (uspio)